

```
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Woliam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
-1- SIMILARITY: Belongs to the ABC transporter family.
CC EMBL; AE009247; AAL43996.1; -; Genomic_DNA.
DR EMBL; AE008366; AAK90206.1; -; Genomic_DNA.
DR PIR; AF2947; AF2947.
DR PIR; D98335; D98335.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome; Nucleotide-binding; Transport.
SQ SEQUENCE 277 AA; 29733 MW; 282D3AA83B0D9124 CRC64;

Query Match 49.5%; Score 51; DB 2; Length 277;
Best Local Similarity 58.8%; Pred. No. 26;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

CY 5 DNAVLEGSIVKVTGANG 21
Db 25 DGAIAAGSLTAVVGANG 41

RESULT 5
Q8KFB5_CHLTF PRELIMINARY; PRT; 361 AA.
ID Q8KFB5_CHLTF PRELIMINARY;
AC Q8KFB5;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Iron(III) ABC transporter, ATP-binding protein, putative.
GN OrderedLocustNames=CT0412;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobaculum.
OX NCBI_TaxID=1097;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
```

```
DR EMBL; AE006470; AAM71658.1; -; Genomic_DNA.
DR TIGR; CT0412; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome; Membrane; Nucleotide-binding;
KW Transport.
SQ SEQUENCE 361 AA; 39746 MW; 7B4BF8619DB69D3D CRC64;

Query Match 49.5%; Score 51; DB 2; Length 361;
Best Local Similarity 55.0%; Pred. No. 33;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

CY 2 AIPDNAVLEGSIVKVTGANG 21
Db 244 AIPESLVLEGHLEKAFSRNG 263

RESULT 6
Q65DJ1_BACLD PRELIMINARY; PRT; 396 AA.
ID Q65DJ1_BACLD PRELIMINARY;
AC Q65DJ1; Q62P10;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Ywbd (SAM (And some other nucleotide) binding motif, PUA domain).
GN Name=ywbd; OrderedLocustNames=BL03857, BL04060;
OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15383718; DOI=10.1159/000079829;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The complete genome sequence of Bacillus licheniformis DSM13, an
RT organism with great industrial potential."
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,
RA Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
RA Ehrlich S.D., Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species."
RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
DR EMBL; AE017333; AAU42873.1; -; Genomic_DNA.
DR EMBL; CP000002; AAU25501.1; -; Genomic_DNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR InterPro; IPR002478; PUA.
DR InterPro; IPR000051; SAM_bd.
DR SMART; SM00359; PUA; 1.
DR PROSITE; PS50890; PUA; 1.
KW Complete proteome.
SQ SEQUENCE 396 AA; 44897 MW; C1FFD1AC5EEFE2AB CRC64;

Query Match 49.5%; Score 51; DB 2; Length 396;
Best Local Similarity 47.1%; Pred. No. 37;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

OY 5 DNAMEGLVAVTGTANG 21
: : |||:::| ||
Db 30 ENGKEGSLIRITDENG 46

RESULT 7

O45000 CAEEL PRELIMINARY; PRT; 1620 AA.
ID O45000 CAEEL PRELIMINARY; PRT; 1620 AA.
AC O45000;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Intestinal neurexin-like protein 1.
GN Name=itx-1; ORFNames=W03D8.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;

RG The C. elegans sequencing consortium;

RT "Genome sequence of the nematode C. elegans: a platform for

RT investigating biology.";

RL Science 282:2012-2018(1998).

DR EMBL; AF043702; AAK21492.3; -; Genomic DNA.

DR Ensembl; W03D8.6; Caenorhabditis elegans.

DR WormBase; WBGene00002174; itx-1.

DR WormPep; W03D8.6; CE32737.

DR InterPro; IPR000742; EGF_2.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR001791; Laminin_G.

DR InterPro; IPR012680; Laminin_G_2.

DR Pfam; PF00008; EGF_2.

DR Pfam; PF02210; Laminin_G_2; 2.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00282; LamG; 3.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS50026; EGF_3; 3.

DR PROSITE; PS50025; LAM_G_DOMAIN; 4.

KW Complete proteome.

DE SEQUENCE 1620 AA; 180492 MW; 6767E4679CBA04D1 CRC64;

Query Match 49.5%; Score 51; DB 2; Length 1620;

Best Local Similarity 78.6%; Pred. No. 1.5e+02;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MAIPDNAMEGLSV 14
| | | | | | | | | | | | | | | |
Db 779 MTIGDNAMEGLSVV 792

RESULT 8

O4HUZ1 GIBZE

ID O4HUZ1 GIBZE PRELIMINARY; PRT; 368 AA.

AC O4HUZ1;

DT 13-SEP-2005 (TREMBlrel. 31, Created)

DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)

DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=FG11217.1;

OS Gibberella zeae PH-1.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

OX NCBI_TaxID=229533;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PH-1;

DR Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,

RA Arachchi H.M., Barna N., Baertien V., Bloom T., Boguslavsky L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Menus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zalnoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACM01000460; EAA75427.1; -; Genomic DNA.
KW Hypothetical protein.

DE SEQUENCE 368 AA; 40538 MW; ADA3963B94946D5 CRC64;

Query Match 49.0%; Score 50.5; DB 2; Length 368;

Best Local Similarity 63.6%; Pred. No. 41;

Matches 14; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

OY 1 MAIPDN-ALVEGLVAVTGTANG 21
| | | | | | | | | | | | | | | |
Db 1 MAIPDNPAVPGKSVLVLTGVNG 22

OY 1 MAIPDN-ALVEGLVAVTGTANG 21
| | | | | | | | | | | | | | | |
Db 1 MAIPDNPAVPGKSVLVLTGVNG 22

OY 1 MAIPDN-ALVEGLVAVTGTANG 21
| | | | | | | | | | | | | | | |
Db 1 MAIPDNPAVPGKSVLVLTGVNG 22

RESULT 9

O9MM39 9SPER

ID O9MM39 9SPER PRELIMINARY; PRT; 412 AA.

AC O9MM39;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE F1 ATPase alpha subunit (Fragment).

GN Name=atpA;

OS Gnetum ula.

OG Mitochondrion.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales.

OX NCBI_TaxID=3383;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=20226063; PubMed=10760278; DOI=10.1073/pnas.97.8.4092;

RA Bowe L.M., Coat G., dePamphilis C.W.;

RT "Phylogeny of seed plants based on all three genomic compartments:

RT extant gymnosperms are monophyletic and Gnetales' closest relatives

RT are conifers.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:4092-4097(2000).

CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.

DR EMBL; AF209109; AAF64667.1; -; Genomic DNA.

DR HSSP; P19483; 1H8E.

DR SMR; Q9MM39; 1-399.

DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO; GO:0045261; C:proton-transporting ATP synthase complex, C. .; IEA.

DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity. .; IEA.

DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. .; IEA.

DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid. .; IEA.

```

DR GO; GO:0015986; P:ATP synthase coupled proton transport; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR005294; ATP_synthf1_alph.
DR InterPro; IPR004100; ATPase_a/b_N.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR Pfam; PF00006; ATP-synt_ab; 1.
DR Pfam; PF02874; ATP-synt_ab_N; 1.
DR TIGRFAMs; TIGR00962; atpA; 1.
DR PROSITE; PS00152; ATPASE ALPHA BETA; 1.
KW ATP synthase; ATP-binding; CF(1); Hydrogen ion transport; Hydrolase;
KW ion transport; Mitochondrion; Nucleotide-binding; Transport.
SQ SEQUENCE 412 AA; 43993 MW; 824873B6ED21327B CRC64;

Query Match 48.5%; Score 50; DB 2; Length 412;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 DNALGSLVKVTGA 19
|::|::|::|::|:
Db 45 DTAIEGSIVKRTGS 59

RESULT 10
Q9T7A3 GNEGN PRELIMINARY; PRT; 418 AA.
ID Q9T7A3_GNEGN
AC Q9T7A3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE ATPase alpha subunit (Fragment).
GN Name=atp1;
OS Gnetum gnemon (Bago).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales;
OX NCBI_TaxID=3382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20052162; PubMed=10586879; DOI=10.1038/46536;
RA Qiu Y.L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,
RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
RT "The earliest angiosperms: evidence from mitochondrial, plastid and
RT nuclear genomes.";
RC Nature 402:404-407(1999).
RL -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
DR EMBL; AF197617; AAF16949.1; -; Genomic_DNA.
DR HSSP; P19483; 1E79.
DR SMR; Q9T7A3; 1-418.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0045261; C:proton-transporting ATP synthase complex, C..; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity..; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota..; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid..; IEA.
DR GO; GO:0015986; P:ATP synthase coupled proton transport; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR005294; ATP_synthf1_alph.
DR InterPro; IPR000793; ATPase_a/b_C.
DR InterPro; IPR004100; ATPase_a/b_N.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR Pfam; PF00006; ATP-synt_ab; 1.
DR Pfam; PF00306; ATP-synt_ab_C; 1.
DR Pfam; PF02874; ATP-synt_ab_N; 1.
DR TIGRFAMs; TIGR00962; atpA; 1.
DR PROSITE; PS00152; ATPASE ALPHA BETA; 1.
KW ATP synthase; ATP-binding; CF(1); Hydrogen ion transport; Hydrolase;
KW ion transport; Mitochondrion; Nucleotide-binding; Transport.

```

```

FT  NON_TER      1      1
FT  NON_TER      418     418
SQ  SEQUENCE      418 AA; 44753 MW; 0FE55A3B5722901 CRC64;

Query Match
Best Local Similarity 48.5%; Score 50; DB 2; Length 418;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      5 DNAVLEGSIVKVTGA 19
      |::|::|::|::|:
Db      44 DTAIEGSLVKRTGS 58

RESULT 11
Q691P6_ORYSA PRELIMINARY; PRT; 451 AA.
AC Q691P6;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Putative aspartic proteinase nepenthesin I.
   Name=OSJNBa0039D04.13;
   Oryza sativa (japonica cultivar-group).
   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
   OC Ehrhartoideae; Oryzeae; Oryza.
   NCBI_TaxID=39947;
   [1]
   NUCLEOTIDE SEQUENCE.
   RA Sasaki T., Matsumoto T., Katayose Y.;
   RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
   clone:OSJNBa0039D04.";
   RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
   DR EMBL; AF006753; BAD32124.1; -; Genomic_DNA.
   DR Gramene; Q691P6; -.
   DR GO; GO:0004194; F:pepsin A activity; IEA.
   DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
   DR InterPro; IPR001461; Peptidase A1.
   DR InterPro; IPR001969; Pept_Asp_AS.
   DR Pfam; PF00026; Asp. 1.
   DR PRINTS; PR00792; PEPsin.
   DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN 1.
SQ  SEQUENCE      451 AA; 46440 MW; E7F24FED78353C5E CRC64;

Query Match
Best Local Similarity 48.5%; Score 50; DB 2; Length 451;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      5 DNAVLEGSIVKVTGAN 20
      |::|::|::|::|:
Db      243 DSPILFGLAKVTGCN 258

RESULT 12
Q6BIS4_DEBHA PRELIMINARY; PRT; 1048 AA.
ID Q6BIS4_DEBHA PRELIMINARY;
AC Q6BIS4;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Similar to CA1657|IPF16022 Candida albicans IPF16022 unknown
   function.
   OrderedLocusNames=DEHA0G08679g;
   OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
   OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
   OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
   NCBI_TaxID=4959;
   [1]
   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP  STRAIN=ATCC 36239 / CBS 767;
RX  PubMed=15229592; DOI=10.1038/nature02579;
RA  Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA  Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talia E.,

```



```
RA Goffard N., Frangeul L., Aigle M., Anthonard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,
RA Hantraye F., Henequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikoleki M., Ozta S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potter S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR382139; CAG90360.1; -; Genomic_DNA.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Complete proteome; Repeat; WD repeat.
SQ SEQUENCE 1048 AA; 117153 MW; 7BAEB602ACF85373 CRC64;

Query Match 48.5%; Score 50; DB 2; Length 1048;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 PDNAVLGSLVKVTGANG 21
DB 26 PDNFVIEGNLMAYTASGG 43

RESULT 13
Q8G456_BIFLO PRELIMINARY; PRT; 3172 AA.
AC Q8G456;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fas.
GN Name=fas; OrderedLocusNames=BL1537;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Scheil M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Erdmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014295; AAN25329.1; -; Genomic_DNA.
DR HSSP; P25715; IMLA.
DR GO; GO:0005835; C:fatty acid synthase complex; IEA.
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR004136; 2npnop_dioxygen.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR003965; Fatty_acid_synth.
DR InterPro; IPR000794; ketoacyl_synth.
DR InterPro; IPR002539; Maoc_dehydratas.
DR Pfam; PF00698; Acyl_transfer_1; 1.
DR Pfam; PF00109; ketoacyl_synth_1.
DR Pfam; PF02801; ketoacyl_synth_C; 1.
DR Pfam; PF01575; Maoc_dehydratas; 1.
DR Pfam; PF03060; NPD; 1.
```

```
DR PRINTS; PR01483; PASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 3172 AA; 338131 MW; 351478B6BF20CFDE CRC64;

Query Match 48.5%; Score 50; DB 2; Length 3172;
Best Local Similarity 50.0%; Pred. No. 4.2e+02;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 AIPDNAVLGSLVKVTGANG 21
DB 2026 AVVDSALDAPAEKVTGSNG 2045

RESULT 14
Q9HYS1_PSEAE PRELIMINARY; PRT; 592 AA.
ID Q9HYS1_PSEAE PRELIMINARY;
AC Q9HYS1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Probable short-chain dehydrogenase.
GN OrderedLocusNames=PA3324;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004754; AAG06712.1; -; Genomic_DNA.
DR PIR; B83231; B83231.
DR HSSP; Q724W1; IPR9.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR InterPro; IPR000639; Epox_hydrolase.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00561; Abhydrolase_1; 1.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00412; EFOXHYDRLASE.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 592 AA; 65770 MW; 7EA0D0996D4D0157 CRC64;

Query Match 47.6%; Score 49; DB 2; Length 592;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 PDNAVLGSLVKVTGANG 21
DB 318 PDASSMNGKLVVTGAGG 335

RESULT 15
Q4ZUUS_PSESY PRELIMINARY; PRT; 732 AA.
ID Q4ZUUS_PSESY PRELIMINARY;
AC Q4ZUUS;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
```

```

DT 13-SEP-2005 (TReMBLrel. 31, last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, last annotation update)
DE Outer membrane autotransporter barrel precursor.
DE ORFNames=Psyr_2034;
OS Pseudomonas syringae pv. syringae B728a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205918;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG DOE Joint Genome Institute;
RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
RA Nolan M., Goltsman E., Thiel J., Malfatti S., Lapidus A., Deter J.C.,
RA Land M., Richardson P.M., Kyriides N.C., Ivanova N.;
RT "Comparison of two complete genome sequences of Pseudomonas syringae
RT pv. syringae B728a and pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Loper J.;
RN Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN (3)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Fell H., Fell W.S., Lindow S.E.;
RN Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RA Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
EMBL, CP000075; AAY37077.1; -; Genomic_DNA.
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR004899; Pectactin.
DR InterPro; IPR003991; Pectactin_C.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF03212; Pectactin; 1.
DR PRINTS; PR01484; PRTACTINFAMILY.
DR TIGRfams; TIGR01414; autotrans_barl; 1.
DR Signal.
FT SIGNAL.
FT SIGNAL.
SQ SEQUENCE 732 AA; 75763 MW; 4C355FC1A478A628 CRC64;
Query Match 47.6%; Score 49; DB 2; Length 732;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 6 NAVLEGLVKVTGANG 21
| | | | | | | | | |
| | | | | | | | | |
DB 148 NLTLEGTSVTATGANG 163

```

Search completed: March 11, 2006, 05:08:44
Job time : 125 secs

This Page Blank (uspto)

DR N-PSDB; AAH74579.
XX
PT Designing synthetic nucleic acid sequences for improved amplification,
PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.
XX
PS Claim 6; Page 66-67; 117pp; English.
XX
CC The present sequence represents a NADPH-dependent aldehyde reductase
CC (AR2). The polynucleotide sequence was modified using the method of
CC the invention. The specification describes a method for designing a
CC synthetic polynucleotide. The method comprises providing a starting
CC polynucleotide, determining the predicted free energy of folding per base
CC of the polynucleotide, modifying the polynucleotide by replacing a codon
CC with a different codon to provide a modified polynucleotide, determining
CC free energy of folding per base of the modified polynucleotides, and
CC comparing this with that of the original polynucleotide. The method is
CC useful for developing nucleic acid sequences that enhance expression of
CC the encoded protein in a heterologous host. The design and preparation of
CC the synthetic genes are used in application of gene shuffling, directed
CC evolution and molecular breeding methods. The method allows expression of
CC genes from various organisms such as mammals, plants, yeast, fungi and
CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
CC hosts at commercially viable levels, in particular proteins with low
CC yield such as methionine gamma-lyase from P. putida. (Updated on 11-SEP-
CC 2003 to standardise OS field)
XX
SQ Sequence 343 AA;

Query Match 66.8%; Score 65.5; DB 4; Length 343;
Best Local Similarity 88.9%; Pred. No. 0.025;
Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4 DNAVL-EGSLVKVTGANG 20
Db 5 DNAVLPEGSLVLTGANG 22
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |

RESULT 5
AAG63563 ID AAG63563 standard; protein; 343 AA.
XX
AC AAG63563;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of a modified NADPH-dependent aldehyde reductase.
XX
KM Methionine gamma-lyase; mdaA gene; free folding energy; gene shuffling;
KM directed evolution; molecular breeding;
KM NADPH-dependent aldehyde reductase.
XX
OS Synthetic.
OS Sporidiobolus salmonicolor.
XX
PN WO200155342-A2.
XX
PD 02-AUG-2001.
XX
PF 31-JAN-2001; 2001WO-US003186.
XX
PR 31-JAN-2000; 2000US-00494921.
PR 08-DEC-2000; 2000US-00734237.
XX
PA (BIOC-) BIOCATALYTICS INC.
XX
PI Rozzell DJ, Bul P, Hua L;
XX WPI, 2001-483235/52.
DR N-PSDB; AAH74580.
XX
PT Designing synthetic nucleic acid sequences for improved amplification,

PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.
XX
PS Claim 6; Page 67-68; 117pp; English.
XX
CC The present sequence represents a modified NADPH-dependent aldehyde
CC reductase (AR2). The polynucleotide sequence was modified using the
CC method of the invention. The specification describes a method for
CC designing a synthetic polynucleotide. The method comprises providing a
CC starting polynucleotide, determining the predicted free energy of folding
CC per base of the polynucleotide, modifying the polynucleotide by replacing
CC a codon with a different codon to provide a modified polynucleotide,
CC determining free energy of folding per base of the modified
CC polynucleotides, and comparing this with that of the original
CC polynucleotide. The method is useful for developing nucleic acid
CC sequences that enhance expression of the encoded protein in a
CC heterologous host. The design and preparation of the synthetic genes are
CC used in application of gene shuffling, directed evolution and molecular
CC breeding methods. The method allows expression of genes from various
CC organisms such as mammals, plants, yeast, fungi and bacteria in
CC prokaryotic hosts, such as Escherichia coli and eukaryotic hosts at
CC commercially viable levels, in particular proteins with low yield such as
CC methionine gamma-lyase from P. putida
XX
SQ Sequence 343 AA;

Query Match 66.8%; Score 65.5; DB 4; Length 343;
Best Local Similarity 88.9%; Pred. No. 0.025;
Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4 DNAVL-EGSLVKVTGANG 20
Db 5 DNAVLPEGSLVLTGANG 22
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |

RESULT 6
ADO56196 ID ADO56196 standard; protein; 343 AA.
XX
AC ADO56196;
XX
DT 12-AUG-2004 (first entry)
XX
DE Sporidiobolus salmonicolor ketoreductase, aldehyde reductase II.
XX
KM (R)-hydroxy ester; alpha-keto ester; ketoreductase; enzyme;
KM aldehyde reductase II.
XX
OS Sporidiobolus salmonicolor.
XX
PN US2004101937-A1.
XX
PD 27-MAY-2004.
XX
PF 08-JUL-2003; 2003US-00616320.
XX
PR 10-JUL-2002; 2002US-0394761P.
XX
PA (MOOR/) MOORE J C.
PA (STUR/) STURR M G.
PA (MCLA/) MCLAUGHLIN K.
PA (KIMJ/) KIM J.
XX
PI Moore JC, Sturr MG, McLaughlin K, Kim J;
XX WPI, 2004-431257/40.
DR
XX
PT preparing (R)-hydroxy ester involves adding alpha-keto ester to mixture
PT comprising ketoreductase enzyme and non-ketoreductase enzyme components.
XX
PS Disclosure; SEQ ID NO 3; 10pp; English.
XX

CC The invention relates to a method of preparing high yields of an (R) -
CC hydroxy ester by reducing an alpha-keto ester using a ketoreductase
CC enzyme. The present sequence represents Sporobolomyces salmonicolor
CC ketoreductase, aldehyde reductase II used in the method of the invention.
XX
SQ Sequence 343 AA;

Query Match 66.8%; Score 65.5; DB 8; Length 343;
Best Local Similarity 88.9%; Pred. No. 0.025; Mismatches 1; Gaps 1;
Matches 16; Conservative 0; Indels 1; Gaps 1;
OY 4 DNAVLEGSIVKVTGANG 20
Db 5 DNAVLEGSIVKVTGANG 22

RESULT 7
ABP66187
ID ABP66187 standard; protein; 3172 AA.
XX
AC ABP66187;
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:931.
XX
KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW anti-diarrheic; antibacterial; inhibitor of Salmonella; detection;
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
PN EPI227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST) SOC PROD NESTLE SA.
XX
DR WPI; 2002-668397/72.

PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT a probe or primer for detecting and/or identifying Bifidobacterium longum
PT in a biological sample.
XX
PS Claim 3; SEQ ID NO 931; 80pp; English.

XX The present invention describes a polynucleotide (I) comprising a
CC sequence of a Bifidobacterium genome selected from the nucleotide
CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
CC least 90% identity or which hybridises with the sequences given in
CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
CC fusion protein, comprising a sequence selected from 1097 sequences given
CC in ABP65258 to ABP6354 ligated in frame to a polynucleotide encoding a
CC heterologous polypeptide. (I) has anti-diarrheic and antibacterial
CC activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition
CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence

CC Listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office
XX
SQ Sequence 3172 AA;

Query Match 51.0%; Score 50; DB 5; Length 3172;
Best Local Similarity 50.0%; Pred. No. 1.5e+02; Mismatches 6; Gaps 0;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
OY 1 AIPDNAVLEGSIVKVTGANG 20
Db 2026 AVVDSALDPAFAEKVTGSNG 2045

RESULT 8
AAR45942
ID AAR45942 standard; protein; 561 AA.
XX
AC AAR45942;
XX
DT 25-MAR-2003 (revised)
DT 22-JUL-1994 (first entry)
XX
DE Pfu thermostable DNA ligase.
XX
KW Ligase chain reaction; LCR; Pyrococcus furiosus; ligation;
KW hyperthermophilic archaeobacterium; catalysis; blunt-end.
XX
OS Pyrococcus furiosus.
XX
PN WO9402615-A1.
XX
PD 03-FEB-1994.
XX
PF 22-JUL-1993; 93WO-US006939.
XX
PR 23-JUL-1992; 92US-00919140.
XX
PA (STRA-) STRATAGENE.
XX
PI Mathur EJ, Marsh EJ, Schoettlin WE,
XX
DR WPI; 1994-048873/06.
DR N-PSDB; AAQ55729.

PT Thermostable DNA ligase from Pyrococcus furiosus - has improved ligation
PT specificity and stability at high temperatures, giving greater efficiency
PT in ligase chain reaction.
XX
PS Claim 3; Page 62-3; 84pp; English.

XX A thermostable DNA ligase from a hyperthermophilic archaeobacterium, e.g.
CC P. furiosus, catalyses template-dependent ligation at temps. of 30-80
CC degrees C, and retains its catalytic ability when subjected to temps. of
CC 85-100 degrees C. The ligase functions effectively in the ligase chain
CC reaction without significant blunt-end ligation. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 561 AA;

Query Match 49.5%; Score 48.5; DB 2; Length 561;
Best Local Similarity 52.6%; Pred. No. 34; Mismatches 4; Indels 1; Gaps 1;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

OY 2 IPDNAVLEGSIVKVTGANG 20
Db 290 IPEKALVEGELVAI-GENG 307

RESULT 9
AAR96967


```

ID AAR96967 standard; protein; 561 AA.
XX
AC AAR96967;
XX
DT 25-MAR-2003 (revised)
DT 26-SEP-1996 (first entry)
XX
DE Pfu DNA ligase.
XX
KW thermostable ligase; Pfu; Pyrococcus furiosus; ligase chain reaction;
KW LCR; template dependent ligation.
XX
OS Pyrococcus furiosus.
XX
PN US5506137-A.
XX
PD 09-APR-1996.
XX
PF 22-JUL-1993; 93US-00096947.
XX
PR 23-JUL-1992; 92US-00919140.
XX
PA (STRA-) STRATAGENE.
XX
PI Schoettlin WE, Mathur EJ, Marsh EJ;
XX
DR WPI; 1996-200280/20.
XX
DR N-PSDB; AAT14926.
XX
PT Plasmid contg. gene for thermostable DNA ligase - useful in ligase chain
PT reactions, stable up to 100 deg. C.
XX
PS Claim 1; Col 29-32; 29pp; English.
XX
CC The present sequence is that of a purified thermostable DNA ligase
CC isolated from a hyperthermophilic marine archaebacterium, Pyrococcus
CC furiosus (Pfu). The Pfu DNA ligase catalyses template dependent ligation
CC at temperatures of about 30-80 deg.C, and substantially retains its
CC catalytic ability when subjected to temperatures of 85-100 deg.C. It has
CC an estimated mol. wt. of 50-70 kDa. (Updated on 25-MAR-2003 to correct PF
CC field.)
XX
SQ Sequence 561 AA;

Query Match 49.5%; Score 48.5; DB 2; Length 561;
Best Local Similarity 52.6%; Pred. No. 34;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 2 IPDNVLEGLSVKVTGANG 20
   ||:|:|:|:|:|:|:|:|
Db 290 IPEKAIVEGELVAI-GENG 307

RESULT 10
ABU37839
ID ABU37839 standard; protein; 636 AA.
XX
AC ABU37839;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #23366.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Neisseria meningitidis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX

```

```

PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA11709.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 65763; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 636 AA;
SQ
Query Match 49.0%; Score 48; DB 6; Length 636;
Best Local Similarity 54.5%; Pred. No. 49;
Matches 12; Conservative 1; Mismatches 5; Indels 4; Gaps 1,
QY 3 PDN----AVLEGLVKVTGANG 20
||| |:||| |
Db 268 PDNRRLMSADIEGKTVMVTGAGC 289
RESULT 11
ID ADX95802 standard; protein; 947 AA.
XX
XX ADX95802;
AC
XX
XX 21-APR-2005 (first entry)
DT
XX
XX Plant full length insert polypeptide seqid 58466.
XX

```

KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 58466; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 947 AA;
QY
Query Match 49.0%; Score 48; DB 8; Length 947;
Best Local Similarity 55.6%; Pred. No. 78;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Db 256 PPNALLSASLLKVAQNG 273
3 PDNAVLEGLVKVTGANG 20
| |:| |:| |:
RESULT 12
ADY22708
ID ADY22708 standard; protein; 965 AA.
XX
AC ADY22708;

XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 70492.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 70492; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 965 AA;
QY
Query Match 49.0%; Score 48; DB 8; Length 965;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Db 274 PPNALLSASLLKVAQNG 291
3 PDNAVLEGLVKVTGANG 20
| |:| |:| |:
RESULT 12
ADY22708
ID ADY22708 standard; protein; 965 AA.
XX
AC ADY22708;

RESULT 13
ABO69084
ID ABO69084 standard; protein; 802 AA.
XX AC ABO69084;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #1259.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI; 2003-615309/58.
XX DR N-PSDB; ABD02655.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 17830; 455bp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 802 AA;
Query Match 48.0%; Score 47; DB 7; Length 802;
Best Local Similarity 55.6%; Pred. No. 94;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 PDNAVEGSLVKTGANG 20
Db 464 PDARSMNGKLVVVTGAGG 481

RESULT 14
ABU39869
ID ABU39869 standard; protein; 6310 AA.
XX AC ABU39869;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by prokaryotic essential gene #25396.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas putida.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Habelbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA43739.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 67793; 1766bp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6310 AA;
Query Match 48.0%; Score 47; DB 6; Length 6310;
Best Local Similarity 37.5%; Pred. No. 1.1e+03;
Matches 12; Conservative 2; Mismatches 6; Indels 12; Gaps 1;

QY 1 AIPDNVLAADGLSVSGTAAGSTIKVYGPNG 4471
Db 4440 ATPDNVLAADGLSVSGTAAGSTIKVYGPNG 4471

Job time : 117.921 secs

RESULT 15
ABG28376
ID ABG28376 standard; protein; 133 AA.
XX
AC ABG28376;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #28367.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS92563.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 58735; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 46.9%; Score 46; DB 4; Length 133;
Best Local Similarity 52.6%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGLVKVTGAN 19
| | | | | | | | | |
| : | | | | | | | |
Db 5 ACPSKATIPGKTVITVGAN 23

TITLE OF INVENTION: Purified Thermostable Pyrococcus
 TITLE OF INVENTION: Furiosus DNA Ligase
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 2730 Sand Hill Road
 CITY: Menlo Park
 STATE: California
 COUNTRY: USA
 ZIP: 94025
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 5.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/916,232
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/919,140
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Albert P. Halluin
 REGISTRATION NUMBER: 25,227
 REFERENCE/DOCKET NUMBER: 8142-013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-854-3660
 TELEFAX: 415-854-3694
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 561 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Pyrococcus Furiosus
 INDIVIDUAL ISOLATE: DSM #3638
 CELL TYPE: unicellular organism
 IMMEDIATE SOURCE:
 LIBRARY: Pyrococcus Furiosus Genomic DNA
 CLONE: pEM1

```

Query Match          49.5%; Score 48.5; DB 2; Length 561;
Best Local Similarity 52.6%; Pred No. 7.1;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

OY      2 IPDNAVLEGS LVKVTGANG 20
      ||:|:|:|:|:|:|:|:|
DB      290 IPEKAIVEGELVAI-GENG 307

RESULT 6
PCT-US93-06939-1
/ Sequence 1, Application PC/TUS9306939
/ GENERAL INFORMATION:
/   APPLICANT: Eric J. Mathur
/   APPLICANT: Edward J. Marsh
/   APPLICANT: Warren E. Schoettlin
/   TITLE OF INVENTION: Purified Thermostable
/   TITLE OF INVENTION: Pyrococcus Furiosus DNA
/   TITLE OF INVENTION: Ligase
/   NUMBER OF SEQUENCES: 17
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: Limbach & Limbach
/   STREET: 2001 Ferry Building
/   CITY: San Francisco
/   STATE: CA
/   COUNTRY: USA
/   ZIP: 94111
/   COMPUTER READABLE FORM:
/

```

```

1 MEDIUM TYPE: Diskette, 3+ inch, 1.4 Mb storage
2
3 COMPUTER: IBM PC Compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: WordPerfect 5.0
6
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: PCT/US93/06939
9 FILING DATE: 19930722
10
11 CLASSIFICATION:
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: 07/919,140
14 FILING DATE: July 23, 1992
15
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Albert P. Halluin
18 REGISTRATION NUMBER: 25,227
19 REFERENCE/DOCKET NUMBER: STRG 20081 USA
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (415) 433-4150
22 TELEFAX: (415) 433-8716
23 TELEX: 278356
24
25 INFORMATION FOR SEQ ID NO: 1:
26
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 561 amino acids
29 TYPE: amino acid
30 TOPOLOGY: linear
31 MOLECULE TYPE: protein
32
33 ORIGINAL SOURCE:
34 ORGANISM: Pyrococcus Furiosus
35 INDIVIDUAL ISOLATE: DSM #3638
36 CELL TYPE: unicellular organism
37 IMMEDIATE SOURCE:
38 LIBRARY: Pyrococcus Furiosus Genomic DNA
39 CLONE: pEM1
40
41 PCT-US93-06939-1

```

Query Match	49.5%;	Score 48.5;	DB 4;	Length 561;
Best Local Similarity	52.6%;	Pred. No. 7.1;		
Matches	10;	Conservative 4;	Mismatches 4;	Indels 1;
				Gaps 1;
QY	2	IPDNAVLGSLVKVTGANG	20	
		: : : : : : : :		
Db	290	IPEKAIVEGELVAI-GENG	307	

```

RESULT 7
US-09-252-991A-17830
; Sequence 17830, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17830
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17830

```

Query Match	48.0%;	Score 47;	DB 2;	Length 802;
Best Local Similarity	55.6%;	Pred. No. 19;		
Matches 10; Conservative		1; Mismatches	7; Indels	0; Gaps
0y	3 PDNAVLEGSIVKVTGANG 20			
	:			
Db	464 PDARSMNGKLIVVTGAGG 481			

RESULT 8

```

US-09-248-796A-16593
; Sequence 16593, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248, 796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074, 725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096, 409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16593
;
; LENGTH: 308
;
; TYPE: PRT
;
; ORGANISM: Candida albicans
US-09-248-796A-16593

```

Query Match	46.9%	Score 46;	DB 2;	Length 308;
Best Local Similarity	43.5%;	Pred. No. 9.4;		
Matches 10; Conservative	5;	Mismatches 4;	Indels 4;	Gaps 1;

```
QY      2 IPDNAV-----LEGSLVKVTGANG 20
          ||||| : ||| : ||| : |
          ||| : ||| : ||| : |
          ||| : ||| : ||| : |
Db      51 IPDNILDAFSLKGKVASVTTGSSG 73
```

RESULT 9

```

US-10-012-231A-116
; Sequence 116, Application US/10012231A
; Patent No. 6924355
;
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
;
; LENGTH: 331
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
US-10-012-231A-116

```

Query Match	46.9%	Score 46;	DB 2;	Length 331;
Best Local Similarity	52.6%	Pred. No. 10;		
Matches 10; Conservative	1;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      1 AIPDNVLEGLVKVTGAN 19
          | | | : | | | | |
Db      29 ACPSKATIPGKTVIVTGAN 47

```

RESULT 10

```

US-10-015-389A-116
; Sequence 116, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-389A-116

```

Query Match	46.9%	Score 46;	DB 2;	Length 331;
Best Local Similarity	52.6%	Pred. No. 10;		
Matches 10; Conservative	1;	Mismatches 8;	Indels 0;	Gaps 0;

QY 1 AIPDNAVLEGSIVKVTGAN 19
| | : | | | | |
Db 29 ACPSKATIPGKTVIVTGAN 47

RESULT 11
US-10-006

```

: Sequence 116, Application US/10006768A
: Patent No. 6936697
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan I.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2830P1C10
: CURRENT APPLICATION NUMBER: US/10/006,768A
: CURRENT FILING DATE: 2002-03-05
: NUMBER OF SEQ ID NOS: 477
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 116
: LENGTH: 331
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-006-768A-116

```

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-778 <KUR>
A/Cross-references: UNIPROT:Q8UGQ3; UNIPARC:UPI00000D19D6; GB:AE007869; PIDN:AAK86791.1;
C/Genetics:
A/Gene: AGR_C_1799
A/Map position: circular chromosome

Query Match 46.6%; Score 48; DB 2; Length 778;
Best Local Similarity 35.0%; Pred. No. 17;
Matches 7; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 2 AIPDNAVLGSLVKVTGANG 21
|:|::|::|::|::|::|
DB 121 ALPSGLADGTWLVLSGSG 140

RESULT 8
D85608
hypothetical protein Z1183 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: D85608; E85659
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: D85608
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-111 <STO>
A/Cross-references: UNIPROT:Q8X9P2; UNIPARC:UPI00000D07CC; GB:AE005174; NID:g12513991; F
A/Experimental source: strain O157:H7, substrain EDL933
A/Accession: E85659
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-111 <ST2>
A/Cross-references: UNIPARC:UPI00000D07CC; GB:AE005174; NID:g12514511; PIDN:AAG55737.1;
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: Z1183; Z1622

Query Match 45.6%; Score 47; DB 2; Length 111;
Best Local Similarity 41.2%; Pred. No. 3;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 PDNAVLGSLVKVTGAN 20
|::|::|::|::|::|
DB 28 PDNAIIQGAISSVGRN 44

RESULT 9
D90218
LSU ribosomal protein L6AB (rpl6AB) [imported] - Sulfolobus solfataricus
C/Species: Sulfolobus solfataricus
C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C/Accession: D90218
R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A/Description: Sulfolobus solfataricus complete genome.
A/Reference number: A99139
A/Accession: D90218
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-181 <KUR>
A/Cross-references: UNIPROT:Q9UX91; UNIPARC:UPI000013416E; GB:AE006641; NID:g13813871; F
C/Genetics:
A/Gene: rpl6AB
C/Superfamily: ribosomal protein L6/L9

Query Match 45.6%; Score 47; DB 2; Length 181;
Best Local Similarity 47.6%; Pred. No. 5.1;
Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 3 IPDNAV--LEGLVKVTGANG 21
|::|::|::|::|::|
DB 12 IPKNVVVDLKGSIIRIKGPKG 32

RESULT 10
H70326
hypothetical protein aq_294 - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C/Accession: H70326
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
v.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: H70326
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-371 <AOQ>
A/Cross-references: UNIPROT:O66644; UNIPARC:UPI00000562FF; GB:AE000682; NID:g2982979; PI
A/Experimental source: strain VFS
C/Genetics:
A/Gene: aq_294

Query Match 45.6%; Score 47; DB 2; Length 371;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAIPDNAVLGSLVKV 16
:|::|::|::|::|
DB 319 LLIPEKAILGGRVKI 334

RESULT 11
A32687
lysyl endopeptidase (EC 3.4.21.50) precursor - Achromobacter lyticus (strain M497-1)
N/Alternate names: Achromobacter lyticus
C/Species: Achromobacter lyticus
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C/Accession: A32687; A32960
R/Ohara, T.; Makino, K.; Shinagawa, H.; Nakata, A.; Norioka, S.; Sakiyama, F.
J. Biol. Chem. 264, 20625-20631, 1989
A/Title: Cloning, nucleotide sequence, and expression of Achromobacter protease I gene.
A/Reference number: A32687; MUID:90062204; PMID:2684982
A/Accession: A32687
A/Molecule type: DNA
A/Residues: 1-653 <OHA>
A/Cross-references: UNIPROT:P15636; UNIPARC:UPI0000125C8B; GB:J05128; NID:g141803; PIDN
R/Tsunasawa, S.; Masaki, T.; Hirose, M.; Soejima, M.; Sakiyama, F.
J. Biol. Chem. 264, 3832-3839, 1989
A/Title: The primary structure and structural characteristics of Achromobacter lyticus p
A/Reference number: A32960; MUID:89139439; PMID:2492988
A/Accession: A32960
A/Molecule type: protein
A/Residues: 206-473 <TSU>
A/Cross-references: UNIPARC:UPI000011061E
C/Comment: This protein specifically hydrolyzes lysyl peptide bonds including the Lys-P
C/Keywords: hydrolase; serine proteinase
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-205/Domain: propeptide #status predicted <PRO>
F/206-473/Product: achromobacter proteinase I #status experimental <MAT>
F/211-421, 217-285, 241-263/Disulfide bonds: #status experimental
F/262, 318, 399/Active site: His, Asp, Ser #status predicted

Query Match 45.1%; Score 46.5; DB 2; Length 653;
Best Local Similarity 52.4%; Pred. No. 24;
Matches 11; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 1 MAIPDNAVLEGS LVKVTGANG 21
:|||||:|:|:|
Db 568 VAIPDNATVE-SPITVSGRTG 587

RESULT 12
T06029
hypothetical protein T28119.100 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06029
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, March 1999
A:Reference number: 215484
A:Accession: T06029
A:Molecule type: DNA
A:Residues: 1-532 <BEV>
A:Cross-references: UNIPROT:Q9T069; UNIPARC:UPI000000ABF4D; EMBL:AL035709; GSPDB:GN00062;
A:Experimental source: cultivar Columbia; BAC clone T28119
C:Genetics:
A:Gene: ATSP:T28119.100
A:Map position: 4

Query Match 44.7%; Score 46; DB 2; Length 532;
Best Local Similarity 52.9%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 PDNAVLEGS LVKVTGAN 20
|:|:|:|:|:|
Db 209 PKNEVLEGSV I KEVSLN 225

RESULT 13
D95957
probable 2-deoxy-D-glucanase 3-dehydrogenase (EC 1.1.1.125) [imported] - Sinorhizobium m
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 05-Oct-2004
C:Accession: D95957
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D95957
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <KUR>
A:Cross-references: UNIPROT:Q92V10; UNIPARC:UPI000000CB6D2; GB:AL591985; PIDN:CAC49324.1;
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaullt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb21348
C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 43.7%; Score 45; DB 2; Length 254;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAIPDNAVLEGS LVKVTGAN 20
|:|:|:|:|:|
Db 1 MSVPSRFSLAGRRVVTGAN 20

RESULT 14

T33939
probable alcohol dehydrogenase (EC 1.1.1.1) [similarity] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33939
R:Becker, M.; Graves, T.; Wilson, C.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid D2063.
A:Reference number: 221442
A:Accession: T33939
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-408 <BEC>
A:Cross-references: UNIPROT:Q9UAT1; UNIPARC:UPI00000762B1; EMBL:AF125951; PIDN:AAD14690.1
A:Experimental source: strain Bristol N2; clone D2063
C:Genetics:
A:Gene: CBSP:D2063.1
A:Map position: 5
A:Introns: 294/3; 359/3; 373/3
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
C:Keywords: NAD; oxidoreductase; zinc
F;47,70,173/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted

Query Match 43.7%; Score 45; DB 2; Length 408;
Best Local Similarity 45.0%; Pred. No. 26;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIPDNAVLEGS LVKVTGANG 21
|:|:|:|:|:|
Db 181 ALKESEVKSQI VAVTGAGG 200

RESULT 15
A69415
conserved hypothetical protein AF1322 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: A69415
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.,
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: A69415
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-501 <KLE>
A:Cross-references: UNIPROT:O28947; UNIPARC:UPI0000056D84; GB:AE001012; GB:AE000782; NID

Query Match 43.7%; Score 45; DB 2; Length 501;
Best Local Similarity 44.4%; Pred. No. 32;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAIPDNAVLEGS LVKVTG 18
|:|:|:|:|:|
Db 218 VAVNDSGVEGRIVEVNG 235

Search completed: March 11, 2006, 05:09:45
Job time : 19.3333 secs

Query Match 46.9%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGLVKVTGAN 19
| | : | | | | |
Db 29 ACPSKATIPGKTIVITGAN 47

RESULT 12

US-10-015-671A-116
; Sequence 116, Application US/10015671A
; Patent No. 6946263
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan 1.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; CURRENT FILING DATE: 2001-12-11
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-671A-116

Query Match 46.9%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGLVKVTGAN 19
| | : | | | | |
Db 29 ACPSKATIPGKTIVITGAN 47

RESULT 13

US-10-015-393A-116
; Sequence 116, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan 1.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C46
; CURRENT APPLICATION NUMBER: US/10/015,393A

; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-393A-116

Query Match 46.9%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGLVKVTGAN 19
| | : | | | | |
Db 29 ACPSKATIPGKTIVITGAN 47

RESULT 14

US-10-011-833A-116
; Sequence 116, Application US/10011833A
; Patent No. 6951920
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan 1.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C22
; CURRENT APPLICATION NUMBER: US/10/011,833A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-011-833A-116

Query Match 46.9%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGLVKVTGAN 19
| | : | | | | |
Db 29 ACPSKATIPGKTIVITGAN 47

RESULT 15

US-10-006-041A-116
; Sequence 116, Application US/10006041A
; Patent No. 6951921
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan 1.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.


```
/ Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thierby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32256
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(388)
; OTHER INFORMATION: Ceres Seq. ID no. 13592517
US-11-096-568A-32256
```

```
Query Match          41.3%; Score 42.5; DB 7; Length 388;
Best Local Similarity 28.9%; Pred. No. 29;
Matches 11; Conservative 4; Mismatches 2; Indels 21; Gaps 1;
```

```
OY      4 PDNAVLEGS LV-----KVTGAN 20
          |||::|||::|||::|||::|||::|||
Db      160 PDVCVVDGSLTATYEGHPFIQLFVKALGKITGAN 197
```

RESULT 8

```
US-11-096-568A-20894
; Sequence 20894, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thierby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20894
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(210)
; OTHER INFORMATION: Ceres Seq. ID no. 12390165
US-11-096-568A-20894
```

```
Query Match          40.8%; Score 42; DB 7; Length 210;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      2 AIPDNAVLEGS LVK 15
          |||::|||::|||::|||::|||
Db      108 ALPDTSVLRGT LQK 121
```

RESULT 9

```
US-11-096-568A-20893
; Sequence 20893, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thierby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20893
; LENGTH: 277
```

```
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(277)
; OTHER INFORMATION: Ceres Seq. ID no. 12390164
US-11-096-568A-20893
```

```
Query Match          40.8%; Score 42; DB 7; Length 277;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      2 AIPDNAVLEGS LVK 15
          |||::|||::|||::|||::|||
Db      175 ALPDTSVLRGT LQK 188
```

RESULT 10

```
US-11-096-568A-1267
; Sequence 1267, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thierby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1267
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(246)
; OTHER INFORMATION: Ceres Seq. ID no. 13624403
US-11-096-568A-1267
```

```
Query Match          40.3%; Score 41.5; DB 7; Length 246;
Best Local Similarity 57.9%; Pred. No. 24;
Matches 11; Conservative 1; Mismatches 6; Indels 1; Gaps 1;
```

```
OY      4 PDNAV-LEGS LVKVTGANG 21
          |||::|||::|||::|||::|||
Db      221 PKRAVL LQGSQEKALGANG 239
```

RESULT 11

```
US-11-096-568A-1266
; Sequence 1266, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thierby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1266
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(283)
; OTHER INFORMATION: Ceres Seq. ID no. 13624402
US-11-096-568A-1266
```

```
Query Match          40.3%; Score 41.5; DB 7; Length 283;
Best Local Similarity 57.9%; Pred. No. 29;
Matches 11; Conservative 1; Mismatches 6; Indels 1; Gaps 1;
```

Qy 4 PDNAV-LEGS LVKVTGANG 21
Db 258 PKRAVLLOGS QEKALGANG 276

```

RESULT 12
US-11-096-568A-1265
; Sequence 1265, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thery
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 1265
; LENGTH: 316
;
; TYPE: PRT
;
; ORGANISM: Zea mays subsp. mays
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(316)
; OTHER INFORMATION: Ceres Seq. ID no. 13624401
US-11-096-568A-1265

```

Query Match	40.3%;	Score 41.5;	DB 7;	Length 316;
Best Local Similarity	57.9%;	Pred. No. 33;		
Matches 11; Conservative	1;	Mismatches 6;	Indels 1;	Gaps 1;

Qy 4 PDNAV-LEGLVKVTGANG 21
| | | : | | | |
Db 291 PKRAVLLOGSQEKALGANG 309

```

RESULT 13
US-11-052-554A-211
; Sequence 211, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 211
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-211

```

Query Match	40.3%;	Score 41.5;	DB 7;	Length 1562;
Best Local Similarity	52.6%;	Pred. No. 2.3e+02;		
Matches 10;	Conservative 3;	Mismatches 5;	Indels 1;	Gaps 1;

QY	4	PDNAV-LEGS	LKVTGANG	21
		: :	:	
Db	1210	PEEAELRQDLV	KITDANG	1228

RESULT 14
US-11-087-099-4990
; Sequence 4990, Application US/11087099
; Publication No. US20060041961A1

```

; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4990
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Brucella melitensis 16M
US-11-087-099-4990

```

Query Match	39.8%	Score 41;	DB 7;	Length 175;
Best Local Similarity	50.0%;	Pred. No. 19;		
Matches	8;	Conservative	4;	Mismatches 4; Indels 0; Gaps 0;

```
QY      1 MAIPDNAVLEGS LVK 16
          ||||::||:|
Db      127 MEIPDNSLVGSPARV 142
```

```

RESULT 15
US-11-096-568A-12932
; Sequence 12932, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12932
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(234)
; OTHER INFORMATION: Ceres Seq. ID no. 14313384
US-11-096-568A-12932

```

Query Match	39.8%	Score 41;	DB 7;	Length 234;
Best Local Similarity	50.0%	Pred. No. 27;		
Matches	8;	Conservative	3;	Mismatches 5;
				Indels 0;
				Gaps 0;

QY 4 PDNAVLEGS LKVTGA 19
|| : || || ||
Db 71 PDGFCIESNVVKVLGA 86

Search completed: March 11, 2006, 05:16:53
Job time : 12.3333 secs


```
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:53896
; CURRENT APPLICATION NUMBER: US/10/989,488A
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/734,237
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Sporidiobolus salmonicolor
; US-10-989-488A-35
```

```
Query Match      66.5%; Score 68.5; DB 5; Length 343;
Best Local Similarity 81.8%; Pred. No. 0.011;
Matches 18; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```
OY      1 MAIPDNAVLEGSIVKVTGANG 21
      || ||||| ||||| |||||
Db      1 MAKIDNAVLEGSIVLVGTGANG 22
```

RESULT 8

```
US-10-989-488A-37
; Sequence 37, Application US/10989488A
; Publication No. US20050196774A1
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:53896
; CURRENT APPLICATION NUMBER: US/10/989,488A
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/734,237
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Sporidiobolus salmonicolor
; OTHER INFORMATION: NADPH-Dependent Aldehyde Reductase 2, having an ala to gly
; OTHER INFORMATION: mutation at amino acid position 2
; US-10-989-488A-37
```

```
Query Match      66.5%; Score 68.5; DB 5; Length 343;
Best Local Similarity 81.8%; Pred. No. 0.011;
Matches 18; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```
OY      1 MAIPDNAVLEGSIVKVTGANG 21
      || ||||| ||||| |||||
Db      1 MAKIDNAVLEGSIVLVGTGANG 22
```

RESULT 9

```
US-10-437-963-171580
; Sequence 171580, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

```
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 171580
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6979C.1.pep
; US-10-437-963-171580
```

```
Query Match      49.5%; Score 51; DB 4; Length 279;
Best Local Similarity 62.5%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      5 DNAVEGSLVYVTGAN 20
      | : | ||| |||||
Db      71 DSPILVGSIVKVTGNN 86
```

RESULT 10

```
US-10-282-122A-65763
; Sequence 65763, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
```

```
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65763
; LENGTH: 636
```


Search completed: March 11, 2006, 05:16:12
Job time : 97.3333 secs

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-67793

```

```

Query Match      45.6%; Score 47; DB 4; Length 6310;
Best Local Similarity 37.5%; Pred. No. 1.3e+03;
Matches 12; Conservative 2; Mismatches 6; Indels 12; Gaps 1;

```

```

OY      2 AIPDNAVLEGLSVKVTGANG 21
Db      4440 ATPDNLILAADGLSVSGTAAGSTIKVYGPNG 4471

```

```

RESULT 15
US-10-450-763-58735
; Sequence 58735, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 58735
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (92)..(104)
; OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
; OTHER INFORMATION: Identified by eMAYRIX, accession number PR00081B, p-value=6.727e-
; OTHER INFORMATION: 11, raw score of 10.38
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (15)..(131)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
; OTHER INFORMATION: accession name adh_short, E-value=9.1e-05, Pfam score of -12.5
US-10-450-763-58735

```

```

Query Match      44.7%; Score 46; DB 5; Length 133;
Best Local Similarity 52.6%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

```

OY      2 AIPDNAVLEGLSVKVTGANG 20
Db      5 ACPSKATIPGKTIVITGAN 23

```

This Page Blank (uspto)


```

; TITLE OF INVENTION: Purified Thermostable Pyrococcus
; TITLE OF INVENTION: Furiusue DNA Ligase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,232
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/919,140
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Albert P. Halluin
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8142-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pyrococcus Furiusue
; INDIVIDUAL ISOLATE: DSM #3638
; CELL TYPE: unicellular organism
; IMMEDIATE SOURCE:
; LIBRARY: Pyrococcus Genomic DNA
; CLONE: pEM1
; US-08-916-232-1

Query Match          47.1%; Score 48.5; DB 2; Length 561;
Best Local Similarity 52.6%; Pred. No. 7.7;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY      3 IPDNAVLEGLVKVTGANG 21
      ||:|:|:|:|:|:|:|:|:|
Db      290 IPEKAIVEGELVALI-GENG 307

RESULT 6
PCT-US93-06939-1
; Sequence 1, Application PC/TUS9306939
; GENERAL INFORMATION:
; APPLICANT: Eric J. Mathur
; APPLICANT: Edward J. Marsh
; APPLICANT: Warren E. Schoettlin
; TITLE OF INVENTION: Purified Thermostable
; TITLE OF INVENTION: Pyrococcus Furiusue DNA
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Diskette, 3+ inch, 1.4 Mb storage
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06939
; FILING DATE: 19930722
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/919,140
; FILING DATE: July 23, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Albert P. Halluin
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: STRG 20081 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 433-4150
; TELEFAX: (415) 433-8716
; TELEX: 278356
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pyrococcus Furiusue
; INDIVIDUAL ISOLATE: DSM #3638
; CELL TYPE: unicellular organism
; IMMEDIATE SOURCE:
; LIBRARY: Pyrococcus Genomic DNA
; CLONE: pEM1
; PCT-US93-06939-1

Query Match          47.1%; Score 48.5; DB 4; Length 561;
Best Local Similarity 52.6%; Pred. No. 7.7;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY      3 IPDNAVLEGLVKVTGANG 21
      ||:|:|:|:|:|:|:|:|:|
Db      290 IPEKAIVEGELVALI-GENG 307

RESULT 7
US-09-252-991A-17830
; Sequence 17830, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17830
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-17830

Query Match          45.6%; Score 47; DB 2; Length 802;
Best Local Similarity 55.6%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      4 PDNAVLEGLVKVTGANG 21
      ||:|:|:|:|:|:|:|:|:|
Db      464 PDARSMNGKLVVVTGAGG 481
```

RESULT 8

US-08-849-602C-27
; Sequence 27, Application US/08849602C
; Patent No. 6171823
; GENERAL INFORMATION:
; APPLICANT: Fabricius, Helle
; APPLICANT: Hastrup, Sven
; TITLE OF INVENTION: A Process of Producing
; TITLE OF INVENTION: Extracellular Proteins in Bacteria
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61718230 No. 6171823disk of No. 6171823th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,602C
; FILING DATE: 02-JUN-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4139.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-849-602C-27

Query Match 45.1%; Score 46.5; DB 2; Length 653;

Best Local Similarity 52.4%; Pred. No. 20;
Matches 11; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MAIPDNVLEGLVKVTGANG 21

Db 568 VAIPDNATVE-SPTIVSGRTG 587

RESULT 9

5248599-2
; Patent No. 5248599
; APPLICANT: SAKIYAMA, FUMIO;NAKATA, ATSUGO
; TITLE OF INVENTION: ACHROMOBACTER PROTEASE I GENE AND GENE
; PRODUCT THEREOF
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/482,266
; FILING DATE: 20-FEB-1990
; SEQ ID NO:2:
; LENGTH: 653
; 5248599-2

Query Match 45.1%; Score 46.5; DB 6; Length 653;

Best Local Similarity 52.4%; Pred. No. 20;
Matches 11; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MAIPDNVLEGLVKVTGANG 21

Db 568 VAIPDNATVE-SPTIVSGRTG 587

RESULT 10

US-09-248-796A-16593
; Sequence 16593, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16593
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-16593

Query Match 44.7%; Score 46; DB 2; Length 308;

Best Local Similarity 43.5%; Pred. No. 10;
Matches 10; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

Qy 3 IPDNV----LEGLVKVTGANG 21

Db 51 IPDNIDAFSLKGVASVTGSSG 73

RESULT 11

US-10-012-231A-116
; Sequence 116, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-012-231A-116

Query Match 44.7%; Score 46; DB 2; Length 331;

Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 AIPDNVLEGLVKVTGAN 20

Db 29 ACPKSKATIPGKTVIVTGAN 47

```
RESULT 12
US-10-015-389A-116
; Sequence 116, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-389A-116

Query Match          44.7%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      2 AIPDNAVLEGLVKVTGAN 20
Db      29 ACPSKATIPGKTVIVTGAN 47

RESULT 13
US-10-006-768A-116
; Sequence 116, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C10
; CURRENT APPLICATION NUMBER: US/10/006,768A
; CURRENT FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 477
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-768A-116
```

```
Query Match          44.7%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      2 AIPDNAVLEGLVKVTGAN 20
Db      29 ACPSKATIPGKTVIVTGAN 47

RESULT 14
US-10-015-671A-116
; Sequence 116, Application US/10015671A
; Patent No. 6946263
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; CURRENT FILING DATE: 2001-12-11
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-671A-116

Query Match          44.7%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      2 AIPDNAVLEGLVKVTGAN 20
Db      29 ACPSKATIPGKTVIVTGAN 47

RESULT 15
US-10-015-393A-116
; Sequence 116, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C46
```

; CURRENT APPLICATION NUMBER: US/10/015,393A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 116
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-393A-116

Query Match 44.7%; Score 46; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 AIPDNAVLEGSIVKVTGAN 20
| | | : | | | | |
Db 29 ACPSKATIPGKTVIVTGAN 47

Search completed: March 11, 2006, 05:11:14
Job time : 28.6667 secs

DR N-PSDB; AAH74579.
XX
PT Designing synthetic nucleic acid sequences for improved amplification,
PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.
XX
PS Claim 6; Page 66-67; 117pp; English.
XX
CC The present sequence represents a NADPH-dependent aldehyde reductase
CC (AR2). The polynucleotide sequence was modified using the method of
CC the invention. The specification describes a method for designing a
CC synthetic polynucleotide. The method comprises providing a starting
CC polynucleotide, determining the predicted free energy of folding per base
CC of the polynucleotide, modifying the polynucleotide by replacing a codon
CC with a different codon to provide a modified polynucleotide, determining
CC free energy of folding per base of the modified polynucleotides, and
CC comparing this with that of the original polynucleotide. The method is
CC useful for developing nucleic acid sequences that enhance expression of
CC the encoded protein in a heterologous host. The design and preparation of
CC the synthetic genes are used in application of gene shuffling, directed
CC evolution and molecular breeding methods. The method allows expression of
CC genes from various organisms such as mammals, plants, yeast, fungi and
CC bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
CC hosts at commercially viable levels, in particular proteins with low
CC yield such as methionine gamma-lyase from P. putida. (Updated on 11-SEP-
CC 2003 to standardise OS field)
XX
SQ Sequence 343 AA;
SQ
Query Match 66.5%; Score 68.5; DB 4; Length 343;
Best Local Similarity 81.8%; Pred. No. 0.0088;
Matches 18; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 MAIPDNAVL-EGSLVKVTGANG 21
Db 1 MAKIDNAVLPEGSLVLTGANG 22
RESULT 5
AAG63563
ID AAG63563 standard; protein; 343 AA.
XX
AC AAG63563;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of a modified NADPH-dependent aldehyde reductase.
XX
KW Methionine gamma-lyase; mdea gene; free folding energy; gene shuffling;
KW directed evolution; molecular breeding;
KW NADPH-dependent aldehyde reductase.
XX
OS Synthetic.
OS Sporidiobolus salmonicolor.
XX
PN WO200155342-A2.
XX
PD 02-AUG-2001.
XX
PF 31-JAN-2001; 2001WO-US003186.
XX
PR 31-JAN-2000; 2000US-00494921.
PR 08-DEC-2000; 2000US-00734237.
XX
PA (BIOC-) BIOCATALYTICS INC.
XX
PI Rozzell DJ, But P, Hua L;
XX WPI; 2001-483235/52.
DR N-PSDB; AAH74580.
XX
PT Designing synthetic nucleic acid sequences for improved amplification,

PT expression in host cell, by comparing free energy of folding of a
PT starting polynucleotide and a modified polynucleotide having a codon
PT replacement.
XX
PS Claim 6; Page 67-68; 117pp; English.
XX
CC The present sequence represents a modified NADPH-dependent aldehyde
CC reductase (AR2). The polynucleotide sequence was modified using the
CC method of the invention. The specification describes a method for
CC designing a synthetic polynucleotide. The method comprises providing a
CC starting polynucleotide, determining the predicted free energy of folding
CC per base of the polynucleotide, modifying the polynucleotide by replacing
CC a codon with a different codon to provide a modified polynucleotide,
CC determining free energy of folding per base of the modified
CC polynucleotides, and comparing this with that of the original
CC polynucleotide. The method is useful for developing nucleic acid
CC sequences that enhance expression of the encoded protein in a
CC heterologous host. The design and preparation of the synthetic genes are
CC used in application of gene shuffling, directed evolution and molecular
CC breeding methods. The method allows expression of genes from various
CC organisms such as mammals, plants, yeast, fungi and bacteria in
CC prokaryotic hosts, such as Escherichia coli and eukaryotic hosts at
CC commercially viable levels, in particular proteins with low yield such as
CC methionine gamma-lyase from P. putida
XX
SQ Sequence 343 AA;
SQ
Query Match 66.5%; Score 68.5; DB 4; Length 343;
Best Local Similarity 81.8%; Pred. No. 0.0088;
Matches 18; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 MAIPDNAVL-EGSLVKVTGANG 21
Db 1 MAKIDNAVLPEGSLVLTGANG 22
RESULT 6
ADO56196
ID ADO56196 standard; protein; 343 AA.
XX
AC ADO56196;
XX
DT 12-AUG-2004 (first entry)
XX
DE Sporobolomyces salmonicolor ketoreductase, aldehyde reductase II.
XX
KW (R)-hydroxy ester; alpha-keto ester; ketoreductase; enzyme;
KW aldehyde reductase II.
XX
OS Sporidiobolus salmonicolor.
XX
PN US2004101937-A1.
XX
PD 27-MAY-2004.
XX
PF 08-JUL-2003; 2003US-00616320.
XX
PR 10-JUL-2002; 2002US-0394761P.
XX
PA (MOOR/) MOORE J C.
PA (STUR/) STURR M G.
PA (MCLA/) MCLAUGHLIN K.
PA (KIMJ/) KIM J.
XX
PI Moore JC, Sturr MG, McLaughlin K, Kim J;
XX WPI; 2004-431257/40.
DR
XX
PT Preparing (R)-hydroxy ester involves adding alpha-keto ester to mixture
PT comprising ketoreductase enzyme and non-ketoreductase enzyme components.
XX
PS Disclosure; SEQ ID NO 3; 10pp; English.
XX

CC The invention relates to a method of preparing high yields of an (R)-
CC hydroxy ester by reducing an alpha-keto ester using a ketoreductase
CC enzyme. The present sequence represents *Sporobolomyces salmonicolor*
CC ketoreductase, aldehyde reductase II used in the method of the invention.
XX
SQ Sequence 343 AA;

Query Match	66.5%;	Score 68.5;	DB 8;	Length 343;
Best Local Similarity	81.8%;	Pred. No. 0.0088;		
Matches 18;	Conservative	0;	Mismatches 3;	Indels 1;
				Gaps 1;

```
Qy      1 MAIPDNAVL-EGSLVKVTGANG 21
        || ||||| ||||| |||||
Db      1 MAKIDNAVLPEGSLVLTGANG 22
```

RESULT 7
ABP66187
ID ABP66187 standard; protein; 3172 AA.

AC ABP66187;

DT 19-NOV-2002 (first entry)

DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:931.

KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW rotavirus; food composition; pharmaceutical composition.

OS *Bifidobacterium longum*.

PN EP1227152-A1.

PD 31-JUL-2002.

PF 30-JAN-2001; 2001EP-00102050.

PR 30-JAN-2001; 2001EP-00102050.

PA (NEST) SOC PROD NESTLE SA.

DR WPI: 2002-668397/72.

Novel polynucleotide comprising *Bifidobacterium* genome sequence useful as a probe or primer for detecting and/or identifying *Bifidobacterium longum* in a biological sample.

PS Claim 3; SEQ ID NO 931; 80bp; English..

The present invention describes a polynucleotide (I) comprising a sequence of a *Bifidobacterium* genome selected from the nucleotide sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at least 90% identity or which hybridises with the sequences given in ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a fusion protein, comprising a sequence selected from 1097 sequences given in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a heterologous polypeptide. (I) has anti-diarrhetic and antibacterial activities, and can be used as an inhibitor of *Salmonella*. (I) (which is a probe) is useful for the detection and/or identification of *Bifidobacterium longum* in a biological sample. A carrier containing the lactic acid bacterium *Bifidobacterium longum* NCC2705 (CNCM I-2618) can be used for preventing and/or treating diarrhoea brought about by pathogenic bacteria and/or rotavirus. The carrier is a food composition selected from milk, yogurt, curd, cheese, fermented milks, milk based fermented products, ice-creams, fermented cereal based products, milk based powders, infant formula, pet food or a pharmaceutical composition selected from tablets, liquid bacterial suspensions, dried oral supplement, wet oral supplement, dry tube feeding or wet tube feeding. (I) is useful in DNA arrays or chips to carry out analysis of the expression of the *Bifidobacterium* gene. ABQ81844 to ABQ81850 represent *Bifidobacterium* related nucleotide sequences given in the Sequence

CC Listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office
XX
SQ Sequence 3172 AA;

Query Match	48.5%	Score 50;	DB 5;	Length 3172;
Best Local Similarity	50.0%;	Pred. No. 1.7e+02;		
Matches 10; Conservative	4;	Mismatches 6;	Indels 0;	Gaps 0;

```
QY      2 AIPDNVLEGSIVKVTGANG 21
        | : | : | : | | | : | |
Db      2026 AVVDSALDFAEAKVTGSNG 2045
```

RESULT 8
AAR45942
ID AAR45942 standard; protein; 561 AA.

AC AAR45942;

DT 25-MAR-2003 (revised)

DE Pfu thermostable DNA ligase.

aa Ligase chain reaction; LCR; *Pyrococcus furiosus*; ligation;
kw KW hyperthermophilic archaeobacterium; catalysis; blunt-end.

Pyrococcus furiosus.

PN W09402615-A1.

PD 03-FEB-1994.

22-JUL-1993; 93WO-US006939.

PR 23-JUL-1992; 92US-00919140.

PA (STRA-) STRATAGENE.

PI Mathur EJ, Marsh EJ, Schoettlin WE;

DR WPI; 1994-048873/06.

Thermosable DNA ligase from *Pyrococcus furiosus* - has improved ligation specificity and stability at high temperatures, giving greater efficiency in ligase chain reaction.

PS Claim 3; Page 62-3; 84pp; English.

A thermostable DNA ligase from a hyperthermophilic archaebacterium, e.g. *P. furiosus*, catalyses template-dependent ligation at temps. of 30-80 degrees C, and retains its catalytic ability when subjected to temps. of 85-100 degrees C. The ligase functions effectively in the ligase chain reaction without significant blunt-end ligation. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 561 AA;

Query Match	47.1%	Score 48.5;	DB 2;	Length 561;
Best Local Similarity	52.6%;	Pred. No. 38;		
Matches 10;	Conservative 4;	Mismatches 4;	Indels 1;	Gaps 1;

Qy 3 IPDNAVLEGS LVKVTGANG 21
|| : || : || : ||
Db 290 IPEKAIVEGELVAI -GENG 307

RESULT 9
AAR96967

ID AAR96967 standard; protein; 561 AA.
XX
AC AAR96967;
XX
DT 25-MAR-2003 (revised)
DT 26-SEP-1996 (first entry)
XX
DE Pfu DNA ligase.
XX
KW thermostable ligase; Pfu; Pyrococcus furiosus; ligase chain reaction;
KW LCR; template dependent ligation.
XX
OS Pyrococcus furiosus.
XX
PN US5506137-A.
XX
PD 09-APR-1996.
XX
PF 22-JUL-1993; 93US-00096947.
XX
PR 23-JUL-1992; 92US-00919140.
XX
PA (STRA-) STRATAGENE.
XX
PI Schoettlin WE, Mathur EJ, Marsh EJ;
XX
DR WPI; 1996-200280/20.
DR N-PSDB; AAT14926.
XX
XX Plasmid contg. gene for thermostable DNA ligase - useful in ligase chain
PT reactions, stable up to 100 deg. C.
XX
PS Claim 1; Col 29-32; 29pp; English.
XX
CC The present sequence is that of a purified thermostable DNA ligase
CC isolated from a hyperthermophilic marine archaebacterium, Pyrococcus
CC furiosus (Pfu). The Pfu DNA ligase catalyses template dependent ligation
CC at temperatures of about 30-80 deg.C, and substantially retains its
CC catalytic ability when subjected to temperatures of 85-100 deg.C. It has
CC an estimated mol. wt. of 50-70 kDa. (Updated on 25-MAR-2003 to correct PF
CC field.)
XX
SQ Sequence 561 AA;

Query Match 47.1%; Score 48.5; DB 2; Length 561;
Best Local Similarity 52.6%; Pred. No. 38;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 IPDNAVLEGS�VKVTGANG 21
||:|:|:|:|:|:|:|:|:|
Db 290 IPEKAIVEGELVAI-GENG 307

RESULT 10
ABU37839
ID ABU37839 standard; protein; 636 AA.
XX
AC ABU37839;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #23366.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Neisseria meningitidis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX

PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA41709.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 65763; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 636 AA;

Query Match 46.6%; Score 48; DB 6; Length 636;
Best Local Similarity 54.5%; Pred. No. 54;
Matches 12; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 4 PDN----AVLEGS�VKVTGANG 21
|||:|:|:|:|:|:|:|:|:|
Db 268 PDNRLMSADIEGKTVMVTGAGG 289

RESULT 11
ADX95802
ID ADX95802 standard; protein; 947 AA.
XX
AC ADX95802;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 58466.
XX

KM	plant protectant; plant growth regulant; gene therapy; plant;
KM	recombinant DNA construct; physical array; plant breeding marker;
KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KM	extreme osmotic condition; pathogen tolerance; pest tolerance;
KM	growth rate; cell cycle pathway; disease resistance;
KM	galactomannan production; lignin production; plant growth regulator;
KM	yield; plant growth; plant development; seed oil; protein yield;
XX	
OS	Unidentified.
XX	
PN	US2004034888-A1.
XX	
PD	19-FEB-2004.
XX	
PF	28-APR-2003; 2003US-00425114.
XX	
PR	06-MAY-1999; 99US-00304517.
PR	05-NOV-2001; 2001US-00985678.
XX	
PA	(LIU/) LIU J.
PA	(ZHOU/) ZHOU Y.
PA	(KOVA/) KOVALIC D K.
PA	(SCRE/) SCREEN S E.
PA	(TABA/) TABASKA J E.
PA	(CAOY/) CAO Y.
XX	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX	
DR	WPI; 2004-180133/17.
XX	
PT	New recombinant DNA construct, useful for improving plant tolerance to
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT	pests, for conferring increased resistance to plant disease, or for
PT	improving yield.
XX	
PS	Claim 1; SEQ ID NO 58466; 15pp; English.
XX	
CC	The invention describes a recombinant DNA construct comprising a
CC	polynucleotide consisting of a sequence encoding an amino acid sequence
CC	available in electronic form from the US patent office at
CC	ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC	of the invention are also useful in physical arrays of molecules and as
CC	plant breeding markers. The recombinant DNA construct is useful for
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in
CC	plant cells by modification of the cell cycle pathway, for conferring
CC	increased resistance to plant disease, for producing galactomannan,
CC	lignin or plant growth regulators, for increasing the rate of homologous
CC	recombination in plants, for improving yield by modification of
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC	or by providing improved plant growth and development under at least one
CC	stress condition or for modifying seed oil or protein yield and/or
CC	content. This is the amino acid sequence of a plant full length insert
CC	polypeptide that can be used in the recombinant DNA construct of the
CC	invention.
XX	
XX	
SQ	Sequence 947 AA;
QY	
Db	
Query Match	46.6%; Score 48; DB 8; Length 947;
Best Local Similarity	55.6%; Pred. No. 86;
Matches 10; Conservative	2; Mismatches 6; Indels 0; Gaps 0;
4 PDNAVLEGLVKVTGANG 21	
256 PPNALLSASLKVAAQONG 273	
RESULT 12	
ID ADY22708	standard; protein; 965 AA.
XX ADY22708;	
AC	

```
XX      DT     21-APR-2005   (first entry)
XX      DE       Plant full length insert polypeptide seqid 70492.
XX
XX      KW    plant protectant; plant growth regulant; gene therapy; plant;
KW      recombinant DNA construct; physical array; plant breeding marker;
KM      cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KM      extreme osmotic condition; pathogen tolerance; pest tolerance;
KM      growth rate; cell cycle pathway; disease resistance;
KM      galactomannan production; lignin production; plant growth regulator;
KM      yield; plant growth; plant development; seed oil; protein yield;
KM      protein content.
XX
XX      OS    Unidentified.
XX
XX      PN    US2004034888-A1.
XX
XX      PD    19-FEB-2004.
XX
XX      PF    28-APR-2003; 2003US-00425114.
XX
PR      06-MAY-1999; 99US-00304517.
PR      05-NOV-2001; 2001US-00985678.
XX
PA      (LIU/) LIU J.
PA      (ZHOU/) ZHOU Y.
PA      (KOVA/) KOVALIC D K.
PA      (SCRE/) SCREE S E.
PA      (TABA/) TABASKA J E.
PA      (CAOY/) CAO Y.
XX
PI      Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
DR      WPI; 2004-180133/17.
XX
PT      New recombinant DNA construct, useful for improving plant tolerance to
PT      cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT      pests, for conferring increased resistance to plant disease, or for
PT      improving yield.
XX
PS      Claim 1; SEQ ID NO 70492; 15pp; English.
CC
XX      The invention describes a recombinant DNA construct comprising a
CC      polynucleotide consisting of a sequence encoding an amino acid sequence
CC      available in electronic form from the US patent office at
CC      ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC      of the invention are also useful in physical arrays of molecules and as
CC      plant breeding markers. The recombinant DNA construct is useful for
CC      improving plant tolerance to cold, heat, drought, herbicides, extreme
CC      osmotic conditions, pathogens or pests, for manipulating growth rate in
CC      plant cells by modification of the cell cycle pathway, for conferring
CC      increased resistance to plant disease, for producing galactomannan,
CC      lignin or plant growth regulators, for increasing the rate of homologous
CC      recombination in plants, for improving yield by modification of
CC      photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC      or by providing improved plant growth and development under at least one
CC      stress condition or for modifying seed oil or protein yield and/or
CC      content. This is the amino acid sequence of a plant full length insert
CC      polypeptide that can be used in the recombinant DNA construct of the
CC      invention.
XX
SQ      Sequence 965 AA;
Query Match          46.6%; Score 48; DB 8; Length 965;
Best Local Similarity 55.6%; Pred. No. 88;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

RESULT 13
ADY65680
ID ADY65680 standard; protein; 272 AA.
XX
AC ADY65680;
XX
DT 02-JUN-2005 (first entry)
XX
DE S. mansoni protein SEQ ID 1098.
XX
KW Schistosomacide; schistosomiasis; psychiatric disorder; vaccine;
diagnosis.
XX
OS Schistosoma mansoni.
XX
PN WO2005023979-A2.
XX
PD 17-MAR-2005.
XX
PF 10-SEP-2004; 2004WO-BR000170.
XX
PR 11-SEP-2003; 2003US-0502277P.
XX
PA (AMPA-) FUNDACAO AMPARO A PESQUISA DO ESTADO.
XX
PI Verjovski-Almeida S, Leite LCC, Farias LP, Miyasato PA, Kawano T;
PI Demarco R, Garcia JCL, Martins EAL, Ho PL, Nascimento ALTO;
PI Dias-Neto E, Setubal JC, Wenck CFM, Madeira AMBN, Rodrigues V;
PI Gargioni C;
XX
DR WPI; 2005-223357/23.
XX
PT New isolated nucleic acid molecule encoding a Schistosoma mansoni
PT protein, useful for as a vaccine or for preventing, diagnosing, or
PT treating Schistosoma mansoni infection.
XX
PS Disclosure; SEQ ID NO 1098; 52pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC Schistosoma mansoni protein, or its portion which is at least 20 amino
CC acids in length. Also included are an expression vector comprising the
CC nucleic acid operably linked to a promoter, a recombinant cell
CC transformed/transformed with the nucleic acid (or expression vector), an
CC immunogenic composition comprising the nucleic acid in combination with a
CC pharmaceutical adjuvant or carrier, an isolated S. mansoni protein
CC comprising an amino acid sequence encoded by the nucleic acid, an
CC immunogenic composition comprising the isolated S. mansoni protein, an
CC isolated antibody which specifically binds to the isolated S. mansoni
CC protein, a hybridoma cell line which produces the isolated antibody, a
CC method for determining a S. mansoni infection in a subject and a computer
CC readable medium having recorded in it a nucleic acid molecule from S.
CC mansoni genome. The nucleic acid molecule is useful for preventing,
CC diagnosing, or treating S. mansoni infection (schistosomiasis). It can
CC also be used as vaccine against S. mansoni. The present sequence is a
CC Schistosoma mansoni protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. Also, SEQ ID 2141-2152 are
CC mentioned in the specification but are not included in the sequence
CC listing.
XX
SQ Sequence 272 AA;

Query Match 45.6%; Score 47; DB 9; Length 272;
Best Local Similarity 47.6%; Pred. No. 29;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAIPDNVLEGLVKVTGANG 21
:|||||:|:|:|:
Db 140 LAIPDLAASKGIVNVSVCG 160

RESULT 14

ABO69084
ID ABO69084 standard; protein; 802 AA.
XX
AC ABO69084;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #1259.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD02655.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 17830; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 802 AA;

Query Match 45.6%; Score 47; DB 7; Length 802;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 PDNAVLEGLVKVTGANG 21
:|||||:|:|:|:
Db 464 PDARSMNGKLVVVTGAGG 481

RESULT 15
ABU39869
ID ABU39869 standard; protein; 6310 AA.
XX
AC ABU39869;
XX
DT 19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #25396.

Search completed: March 11, 2006, 05:02:25

Job time : 120.667 secs

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Pseudomonas putida.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA43739.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 67793; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 6310 AA;

Query Match 45.6%; Score 47; DB 6; Length 6310;

Best Local Similarity 37.5%; Pred. No. 1.2e+03;
Matches 12; Conservative 2; Mismatches 6; Indels 12; Gaps 1;

OY 2 AIPDNAVL-----EGSLVKVTGANG 21

Db 4440 ATPDNLILAADGLSVSGTAAGSTIKVYGPNG 4471


```
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:53896
; CURRENT APPLICATION NUMBER: US/10/989,488A
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/734,237
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Sporidiobolus salmonicolor
; US-10-989-488A-35
```

```
Query Match      66.8%; Score 65.5; DB 5; Length 343;
Best Local Similarity 88.9%; Pred. No. 0.031;
Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
OY      4 DNAVLEGSGLVKTGANG 20
        ||||| ||||| |||||
Db      5 DNAVLEGSGLVLTGANG 22
```

RESULT 8

```
; Sequence 37, Application US/10989488A
; Publication No. US20050196774A1
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:53896
; CURRENT APPLICATION NUMBER: US/10/989,488A
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/734,237
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Sporidiobolus salmonicolor
; OTHER INFORMATION: NADPH-Dependent Aldehyde Reductase 2, having an ala to gly
; OTHER INFORMATION: mutation at amino acid position 2
; US-10-989-488A-37
```

```
Query Match      66.8%; Score 65.5; DB 5; Length 343;
Best Local Similarity 88.9%; Pred. No. 0.031;
Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
OY      4 DNAVLEGSGLVKTGANG 20
        ||||| ||||| |||||
Db      5 DNAVLEGSGLVLTGANG 22
```

RESULT 9

```
; Sequence 171580, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

```
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 171580
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6979C.1.pep
; US-10-437-963-171580
```

```
Query Match      52.0%; Score 51; DB 4; Length 279;
Best Local Similarity 62.5%; Pred. No. 6.4;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      4 DNAVLEGSGLVKTGAN 19
        ||:| ||||| |||||
Db      71 DSPILVGLAKVTGNN 86
```

RESULT 10

```
; Sequence 65763, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
```

```
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65763
; LENGTH: 636
```

```

; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-65763

Query Match
Best Local Similarity 49.0%; Score 48; DB 4; Length 636;
Matches 12; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 3 PDNAVLEGSIVKVTGANG 20
   ||| :||| |||||
Db 268 PDNRRLMSADIEGKTVMTGAGG 289

RESULT 11
US-10-425-115-273937
; Sequence 273937, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 273937
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(892)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_181416C.1.pep
US-10-425-115-273937

Query Match
Best Local Similarity 49.0%; Score 48; DB 4; Length 892;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 PDNAVLEGSIVKVTGANG 20
   ||| :||| |||||
Db 202 PPNALLSASLKVAAQONG 219

RESULT 12
US-10-425-114-58466
; Sequence 58466, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58466
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700383178_FLI.pep
US-10-425-114-58466
```

```

Query Match
Best Local Similarity 49.0%; Score 48; DB 4; Length 947;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 PDNAVLEGSIVKVTGANG 20
   ||| :||| |||||
Db 256 PPNALLSASLKVAAQONG 273

RESULT 13
US-10-425-114-70492
; Sequence 70492, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70492
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73025E08_FLI.pep
US-10-425-114-70492

Query Match
Best Local Similarity 49.0%; Score 48; DB 4; Length 965;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 PDNAVLEGSIVKVTGANG 20
   ||| :||| |||||
Db 274 PPNALLSASLKVAAQONG 291

RESULT 14
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyckind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
```

Search completed: March 11, 2006, 05:16:11
Job time : 93.6984 secs

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67793
;
; LENGTH: 6310
;
; TYPE: PRT
;
; ORGANISM: Pseudomonas putida
US-10-282-122A-67793

```

Query Match	48.0%;	Score 47;	DB 4;	Length 6310;
Best Local Similarity	37.5%;	Pred. No. 1.2e+03;		
Matches 12; Conservative	2;	Mismatches 6;	Indels 12;	Gaps 1;

```

Qy      1 AIPDAVL-----EGSLVKVTGANG 20
      ||| : |||
Db      4440 ATPDNLILADGLSVSGTAEAGSTIKVGPNG 4471

```

RESULT 15
US-10-450-763-58735

```

1 Sequence 58735, Application US/10450763
2 Publication No. US20050196754A1
3 GENERAL INFORMATION:
4 APPLICANT: Hyseq, Inc
5 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
6 FILE REFERENCE: 790CIP3/US
7 CURRENT APPLICATION NUMBER: US/10/450,763
8 PRIOR FILING DATE: 2003-06-11
9 PRIOR APPLICATION NUMBER: PCT/US01/08631
10 PRIOR FILING DATE: 2001-03-30
11 PRIOR APPLICATION NUMBER: 09/540,217
12 PRIOR FILING DATE: 2000-03-31
13 PRIOR APPLICATION NUMBER: 09/649,167
14 PRIOR FILING DATE: 2000-08-23
15 NUMBER OF SEQ ID NOS: 60736
16 SOFTWARE: Custom
17 SEQ ID NO 58735
18 LENGTH: 133
19 TYPE: PRT
20 ORGANISM: Homo sapiens
21 FEATURE:
22 NAME/KEY: DOMAIN
23 LOCATION: (92)..(104)
24 OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
25 OTHER INFORMATION: Identified by EMATRIX, accession number PR00081B, p-value=6.727e
26 OTHER INFORMATION: 11, raw score of 10.38
27 FEATURE:
28 NAME/KEY: DOMAIN
29 LOCATION: (15)..(131)
30 OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
31 OTHER INFORMATION: accession name adh_short, E-value=9.1e-05, Pfam score of -12.5
32 US-10-450-763-58735

```

Query Match	46.9%;	Score 46;	DB 5;	Length 133;
Best Local Similarity	52.6%;	Pred. No. 18;		
Matches 10;	Conservative 1;	Mismatches 8;	Indels 0;	Gaps 0;

QY 1 AIPDNVLEGLVKVTGAN 19
| | : | | | | |
Db 5 ACPKATIPGKTVIVTGAN 23

This Page Blank (uspto)

Db 1 MAQVDNEV--KGKVLITGASG 20

RESULT 4

Q6MYZ7 ASPFU PRELIMINARY; PRT; 352 AA.

AC Q6MYZ7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Oxidoreductase, putative (EC 1.-.-.-).
GN ORFNames=AfA24A6.070C;
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;

NUCLEOTIDE SEQUENCE.

RP PubMed=14998527; DOI=10.1016/j.fgb.2003.12.003;
RX Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,
RA Foster N., Fraser A., Harris D., Larke N., Murphy L., Humphray S.,
RA O'Neill S., Perte M., Price C., Rabinowitsch E., Rajandream M.-A.,
RA Salzberg S., Saunders D., Seegar K., Sharp S., Warren T.,
RA Denning D.W., Barrell B., Hall N.;
RT "Insight into the genome of Aspergillus fumigatus: analysis of a 922
kb region encompassing the nitrate assimilation gene cluster.";
RL Fungal Genet. Biol. 41:443-453(2004).
DR EMBL; BX649605; CAB47862.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
KW Oxidoreductase.

SEQUENCE 352 AA; 39420 MW; 0E84223EEC87ABF9 CRC64;

Query Match 54.2%; Score 58; DB 2; Length 352;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 MAKIDNAVLPEGSLVLTGANG 22
Db 1 MGSIAPTIVPPGGLVLTGVNG 22

RESULT 5
Q4MSW9 ASPFU PRELIMINARY; PRT; 352 AA.

AC Q4MSW9;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Aldehyde reductase II.
GN ORFNames=Afu1g11360;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;

NUCLEOTIDE SEQUENCE.

RP STRAIN=Af293;
RC Nierman W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley J.,
RA Arroya J., Bertram M., Abe K., Archer D.B., Berneto C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khoult H., Kitamoto K., Kobayashi T., Kulikarni R.,
RA Kumagai T., Latton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majors W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penava M.A., Perte M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,

RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Roming C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Abai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
Aspergillus fumigatus.";
RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHF01000004; EAL90463.1; -; Genomic DNA.
SQ SEQUENCE 352 AA; 39420 MW; 0E84223EEC87ABF9 CRC64;

Query Match 54.2%; Score 58; DB 2; Length 352;
Best Local Similarity 50.0%; Pred. No. 5;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 MAKIDNAVLPEGSLVLTGANG 22
Db 1 MGSIAPTIVPPGGLVLTGVNG 22

RESULT 6
Q5ASZ7 EMENI PRELIMINARY; PRT; 341 AA.

AC Q5ASZ7;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein.

GN ORFNames=AN8583.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;

NUCLEOTIDE SEQUENCE.

RP STRAIN=FGSC A4;
RC Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Maucelli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RT Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000158; EAA60617.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 341 AA; 37852 MW; 9381BB9083B3F9EA CRC64;

Query Match 51.4%; Score 55; DB 2; Length 341;
Best Local Similarity 66.7%; Pred. No. 13;

Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 8 VLPEGSGLVLTGANG 22
|:|:|||||
Db 7 VIPVGATILVTGANG 21

RESULT 7

Q4KK52_PSEF5
ID Q4KK52_PSEF5 PRELIMINARY; PRT; 398 AA.
AC Q4KK52;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PFL_0235;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodí D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomashow L., Loper J.,
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5."
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL; CP000076; AAY95646.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 398 AA; 43691 MW; B8598CE348B57CBB CRC64;

Query Match 51.4%; Score 55; DB 2; Length 398;
Best Local Similarity 63.2%; Pred. No. 16;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPEGSGLVLTGANG 22
|:|:|||||
Db 13 INKLVIPLGRNLVTGANG 31

RESULT 8

Q4KBV7_PSEF5
ID Q4KBV7_PSEF5 PRELIMINARY; PRT; 684 AA.
AC Q4KBV7;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Phenylacetic acid degradation protein Paan.
GN Name=paan; ORFNames=PFL_3140;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodí D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomashow L., Loper J.,
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5."
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL; CP000076; AAY92410.1; -; Genomic_DNA.
SQ SEQUENCE 684 AA; 73304 MW; 0EB67175E246A23D CRC64;

Query Match 51.4%; Score 55; DB 2; Length 684;
Best Local Similarity 61.1%; Pred. No. 27;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 DNAVLPEGSGLVLTGANG 22
:::|||||
Db 201 ESGILPEGSGLQLVIGATG 218

RESULT 9

Q5KRL6_CORGL
ID Q5KRL6_CORGL PRELIMINARY; PRT; 698 AA.
AC Q5KRL6;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Putative aldehyde dehydrogenase.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=R;
RA Suzuki N., Okayama S., Nonaka H., Tsuge Y., Inui M., Yukawa H.;
RT "Large-Scale Engineering of the Corynebacterium glutamicum Genome."
RL Appl. Environ. Microbiol. 71:3369-3372(2005).
DR EMBL; AB193033; BAD84061.1; -; Genomic_DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR002539; Maoc dehydratase.
DR InterPro; IPR011966; Paan-DH.
DR Pfam; PF00171; Aldehyd; 1.
DR TIGRFAMs; TIGR02278; Paan-DH; 1.
SQ SEQUENCE 698 AA; 75061 MW; E9781F184A0D8B57 CRC64;

Query Match 51.4%; Score 55; DB 2; Length 698;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 4 IDNAVLPEGSGLVLTGANG 22
:::|||||
Db 204 VESGILPEGSGLQISGSVG 222

RESULT 10

Q9RRV4_DEIRA
ID Q9RRV4_DEIRA PRELIMINARY; PRT; 700 AA.
AC Q9RRV4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Aldehyde dehydrogenase.
GN OrderedlocusNames=DR2381;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Uitterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";

```
RL Science 286:1571-1577(1999).
DR EMBL; AE002069; AAF11927.1; -; Genomic_DNA.
DR PIR; F75279; F75279.
DR TIGR; DR2381; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR InterPro; IPR002539; Maoc_dehydratas.
DR InterPro; IPR011966; Paan-DH.
DR Pfam; PF00171; Aldehyd; 1.
DR Pfam; PF01575; Maoc_dehydratas; 1.
DR TIGRFAMs; TIGR02278; Paan-DH; 1.
KW Complete proteome.
SQ SEQUENCE 700 AA; 75493 MW; E3F6FDA472C9E932 CRC64;

Query Match          51.4%; Score 55; DB 2; Length 700;
Best Local Similarity 63.2%; Pred. No. 28;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 IDNAVLPEGSLVLTGAN 22
Db 203 IASGILPEGSLQLVTGDPG 221

RESULT 11
Q6WE58 PINMO
ID Q6WE58_PINMO PRELIMINARY; PRT; 206 AA.
AC Q6WE58;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative NBS-LRR protein C618 (Fragment).
GN Name=RGA;
OS Pinus monticola (Western white pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Strobus.
OX NCBI_TaxID=3345;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14586641; DOI=10.1007/s00438-003-0940-1;
RA Liu J.-J., Ekramoddoullah A.K.M.;
RT "Isolation, genetic variation and expression of TIR-NBS-LRR resistance
RT gene analogs from western white pine (Pinus monticola Dougl. ex. D.
RT Don.).";
RL Mol. Genet. Genomics 270:432-441(2004).
DR EMBL; AY294116; AAQ57161.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 206 AA; 23705 MW; 9A72894D0F7CECD9 CRC64;

Query Match          50.5%; Score 54; DB 2; Length 206;
Best Local Similarity 68.8%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGAN 21
Db 20 NDVLPESGSLILVTSRN 35

RESULT 12
Q6WE60 PINMO
ID Q6WE60_PINMO PRELIMINARY; PRT; 206 AA.
AC Q6WE60;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative NBS-LRR protein C605 (Fragment).
GN Name=RGA;
OS Pinus monticola (Western white pine).
```

```
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Strobus.
OX NCBI_TaxID=3345;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14586641; DOI=10.1007/s00438-003-0940-1;
RA Liu J.-J., Ekramoddoullah A.K.M.;
RT "Isolation, genetic variation and expression of TIR-NBS-LRR resistance
RT gene analogs from western white pine (Pinus monticola Dougl. ex. D.
RT Don.).";
RL Mol. Genet. Genomics 270:432-441(2004).
DR EMBL; AY294107; AAQ57159.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 206 AA; 23720 MW; C2D2A34E5A1FBAEC CRC64;

Query Match          50.5%; Score 54; DB 2; Length 206;
Best Local Similarity 68.8%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGAN 21
Db 20 NDVLPESGSLILVTSRN 35

RESULT 13
Q6WE61 PINMO
ID Q6WE61_PINMO PRELIMINARY; PRT; 206 AA.
AC Q6WE61;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative NBS-LRR protein C603 (Fragment).
GN Name=RGA;
OS Pinus monticola (Western white pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Strobus.
OX NCBI_TaxID=3345;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14586641; DOI=10.1007/s00438-003-0940-1;
RA Liu J.-J., Ekramoddoullah A.K.M.;
RT "Isolation, genetic variation and expression of TIR-NBS-LRR resistance
RT gene analogs from western white pine (Pinus monticola Dougl. ex. D.
RT Don.).";
RL Mol. Genet. Genomics 270:432-441(2004).
DR EMBL; AY294106; AAQ57158.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 206 AA; 23750 MW; DECF629247DB19F8 CRC64;

Query Match          50.5%; Score 54; DB 2; Length 206;
Best Local Similarity 68.8%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGAN 21
Db 20 NDVLPESGSLILVTSRN 35

RESULT 14
Q6WE65 PINMO
ID Q6WE65_PINMO PRELIMINARY; PRT; 206 AA.
AC Q6WE65;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
```

Search completed: March 11, 2006, 05:08:43
Job time : 132.905 secs

```
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Putative NBS-LRR protein GM01 (Fragment).
GN Name=RGA;
OS Pinus monticola (Western white pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Strobus.
OX NCBI_TaxID=3345;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14586641; DOI=10.1007/s00438-003-0940-1;
RA Liu J.-J., Ekramoddoullah A.K.M.;
RT "Isolation, genetic variation and expression of TIR-NBS-LRR resistance
RT gene analogs from western white pine (Pinus monticola Dougl. ex. D.
RT Don.).";
RL Mol. Genet. Genomics 270:432-441(2004).
DR EMBL; AY294096; AAQ57154.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 206 AA; 23808 MW; 73A5FFCDB9F5C0D7 CRC64;

Query Match 50.5%; Score 54; DB 2; Length 206;
Best Local Similarity 68.8%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGAN 21
DB 20 NDVLPVSGSLVLTSRN 35

RESULT 15
Q6HC40_BACHK PRELIMINARY; PRT; 250 AA.
ID Q6HC40;
AC Q6HC40;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Glucanase permease, C-terminal, GntP family.
GN Name=gntP; OrderedLocustNames=BT9727.4574;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=180856;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27.";
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE017355; AAT60986.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015128; F:glucanase transporter activity; IEA.
DR GO; GO:0015725; P:glucanase transport; IEA.
DR InterPro; IPR003474; Gln_transporter.
DR Pfam; PF02447; GntP_permease; 1.
KW Complete proteome.
SQ SEQUENCE 250 AA; 26921 MW; 3BC16A20C586C23C CRC64;

Query Match 50.5%; Score 54; DB 2; Length 250;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 5 DNAVLPESGLVLTGAN 22
DB 104 DESLPPVGSILITIGAG 121
```

C/Accession: A83785
R/Takami, H.; Nakabone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: A83785
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-274 <STO>
A/Cross-references: UNIPROT:Q9KDY0; UNIPARC:UPI00000C3AAF; GB:AP001510; GB:BA000004; NID
A/Experimental source: strain C-125
C/Genetics:
A/Gene: fhuc
C/Superfamily: inner membrane protein malK; ATP-binding cassette homology

Query Match 46.7%; Score 50; DB 2; Length 274;
Best Local Similarity 47.1%; Pred. No. 6.4;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGANG 22
| : | | | : | : | | |
DB 23 NIAIPEGKITITGANG 39

RESULT 8
AIO445
probable zinc-binding dehydrogenase [imported] - *Yersinia pestis* (strain CO92)
C/Species: *Yersinia pestis*
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C/Accession: AIO445
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AIO445
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-325 <KUR>
A/Cross-references: UNIPROT:Q8ZAW6; UNIPARC:UPI00000DC7DA; GB:AL590842; PIDN:CAC93133.1;
C/Genetics:
A/Gene: YPO3663
C/Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

Query Match 45.8%; Score 49; DB 2; Length 325;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 MAKIDNAVLPESGLVLTGANG 22
| | | | | : | : | | | : |
DB 136 MALEDGCVTPKSGEVVVTGASG 157

RESULT 9
G71154
probable adenylosuccinate synthetase - *Pyrococcus horikoshii*
C/Species: *Pyrococcus horikoshii*
C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C/Accession: G71154
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushiida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A/Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: G71154
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-339 <KAW>
A/Cross-references: UNIPROT:O58187; UNIPARC:UPI0000132B05; GB:AP000002; NID:G3236129; PI
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:

A/Gene: PH0438

Query Match 45.8%; Score 49; DB 2; Length 339;
Best Local Similarity 64.7%; Pred. No. 11;
Matches 11; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGANG 22
| | | | | : | : | | | : |
DB 164 NDALDEGSLVLTGTOG 180

RESULT 10
E75385
carbohydrate kinase, PKB family - *Deinococcus radiodurans* (strain R1)
C/Species: *Deinococcus radiodurans*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: E75385
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; H
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mah
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: E75385
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-311 <WHI>
A/Cross-references: UNIPROT:Q9RU69; UNIPARC:UPI00000D3E78; GB:AE001996; GB:AE000513; NID
C/Genetics:
A/Gene: DRI525
A/Map position: 1
C/Superfamily: ribokinase

Query Match 44.9%; Score 48; DB 2; Length 311;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 6 NAVLPEGSLVLTGGA 20
| : | | | | : | : | | | : |
DB 242 DAVLPEGTVDITGA 256

RESULT 11
B90054
glucanate permease [imported] - *Staphylococcus aureus* (strain N315)
C/Species: *Staphylococcus aureus*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: B90054
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: B90054
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-452 <KUR>
A/Cross-references: UNIPROT:Q99RD0; UNIPARC:UPI000005488A; GB:BA000018; PID:gl3702455; P
A/Experimental source: strain N315
C/Genetics:
A/Gene: gntP
C/Superfamily: D-serine permease

Query Match 44.9%; Score 48; DB 2; Length 452;
Best Local Similarity 36.4%; Pred. No. 22;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 1 MAKIDNAVLPESGLVLTGANG 22
| : | | | : | : | : | : |
DB 303 MESVTNAIYPIGMMLLITGGG 324

RESULT 12
F64889
membrane protein maoc - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: F64889
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F64889
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-681 <BLAT>
A:Cross-references: UNIPROT:P77455; UNIPARC:UPI000012EBA1; GB:AE000236; GB:U00096; NID:9
A:Experimental source: strain K-12, substrain MG1655
F:301-317/Domain: transmembrane #status predicted <TM01>
F:588-604/Domain: transmembrane #status predicted <TM02>

Query Match 44.9%; Score 48; DB 2; Length 681;
Best Local Similarity 36.8%; Pred. No. 34;
Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 4 IDNAVLPEGSLVLTGANG 22
:|::|||::|::|::|
DB 198 VDSGLVPEGALSLIGSAG 216

RESULT 13
AE2904
hypothetical protein Atu2670 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AE2904
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
sner, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2904
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: UNIPROT:Q8UC28; UNIPARC:UPI00000D1FA7; GB:AE008688; PIDN:AAI43651.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2670
A:Map position: circular chromosome

Query Match 44.4%; Score 47.5; DB 2; Length 264;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 11; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MAKIDNAVLPEGSLVLTGANG 22
:|||::|||::|::|::|
DB 22 LAKID-LTIEDGSFVTVIGSNG 42

RESULT 14
G97679
probable ATP-binding component of ABC transporter (PA3838) [imported] - Agrobacterium tu
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97679
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughity, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97679
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: UNIPROT:Q8UC28; UNIPARC:UPI00000D1FA7; GB:AE007869; PIDN:AAK88392.1;
C:Genetics:
A:Gene: AGR_C_4841
A:Map position: circular chromosome

Query Match 44.4%; Score 47.5; DB 2; Length 264;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 11; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 MAKIDNAVLPEGSLVLTGANG 22
:|||::|||::|::|::|
DB 22 LAKID-LTIEDGSFVTVIGSNG 42

RESULT 15
T38115
probable ATP-dependent transporter - fission yeast (Schizosaccharomyces pombe) (fragment)
C:Species: Schizosaccharomyces pombe
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: T38115
R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z21771
A:Accession: T38115
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-255 <CON>
A:Cross-references: UNIPROT:O13888; UNIPARC:UPI000006AE18; EMBL:Z98600; PIDN:CAB11251.1;
A:Experimental source: strain 972h-; cosmid c20G4
C:Genetics:
A:Gene: SPDB:SPAC20G4.01
A:Map position: 1
C:Superfamily: yeast hypothetical protein YFL028c; ATP-binding cassette homology

Query Match 43.9%; Score 47; DB 2; Length 255;
Best Local Similarity 64.3%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 LPEGSLVLTGANG 22
||::|||::|||::|
DB 2 LPKGSRITLLVGANG 15

Search completed: March 11, 2006, 05:09:45
Job time : 22.254 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 05:02:47 ; Search time 18.4127 Seconds
(without alignments)
104.511 Million cell updates/sec

Title: US-10-616-320A-1
Perfect score: 98
Sequence: 1 AIPDNAVLGSLVKVTGANG 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	52.0	277	2 AF2947	hypothetical prote
2	51	52.0	277	2 D98335	ABC transporter, A
3	49	50.0	592	2 B83231	probable short-cha
4	48	49.0	636	2 G81983	pilin glycosylatio
5	48	49.0	778	2 AF2697	two component sens
6	48	49.0	778	2 F97479	non-motile and pha
7	47	48.0	111	2 D85608	hypothetical prote
8	47	48.0	181	2 D90218	1SU ribosomal prot
9	46	46.9	371	2 H70326	hypothetical prote
10	46	46.9	532	2 T06029	hypothetical prote
11	46	46.9	1647	2 T32934	hypothetical prote
12	45.5	46.4	653	2 A32687	lysyl endopeptidas
13	45	45.9	408	2 T33939	probable alcohol d
14	45	45.9	536	2 B46230	RNA-binding protei
15	45	45.9	615	2 S24761	Cpo 61.1 protein -
16	45	45.9	616	2 A46230	RNA-binding protei
17	45	45.9	930	2 T08588	hypothetical prote
18	44	44.9	327	4 S29876	probable H+-transp
19	44	44.9	491	2 S17916	H+-transporting tw
20	44	44.9	501	2 A69415	conserved hypothet
21	44	44.9	506	2 S33922	H+-transporting tw
22	44	44.9	506	2 S46508	H+-transporting tw
23	44	44.9	507	1 PWRPA	H+-transporting tw
24	44	44.9	507	2 S12309	H+-transporting tw
25	44	44.9	508	1 PWSMAM	H+-transporting tw
26	44	44.9	508	2 S29792	H+-transporting tw
27	44	44.9	508	2 S26979	H+-transporting tw
28	44	44.9	509	1 PMNTAC	H+-transporting tw
29	44	44.9	509	1 PWRZAM	H+-transporting tw

30	44	44.9	509	1 PWTAM	H+-transporting tw
31	44	44.9	510	2 S19261	H+-transporting tw
32	44	44.9	510	2 S10997	H+-transporting tw
33	44	44.9	511	2 S07316	H+-transporting tw
34	43	43.9	87	1 F71194	probable ribosomal
35	43	43.9	206	2 F69490	conserved hypothet
36	43	43.9	232	2 G84382	cobalt transport A
37	43	43.9	240	2 AB2750	hypothetical prote
38	43	43.9	240	2 A97531	branched-chain ami
39	43	43.9	252	2 AH3618	high-affinity bran
40	43	43.9	283	2 S41412	tetrahydroxynaphth
41	43	43.9	346	2 C81654	tryptophanyl-tRNA
42	43	43.9	362	1 LPECPA	rare lipoprotein A
43	43	43.9	362	2 C85563	a minor lipoprotei
44	43	43.9	362	2 G90712	a minor lipoprotei
45	43	43.9	425	2 H70456	modulation competi

ALIGNMENTS

RESULT 1
AF2947
hypothetical protein Atu3180 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AF2947
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AF2947
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-277 <KUR>
A/Cross-references: UNIPROT:Q8UB37; UNIPARC:UPI00000D20E7; GB:AE008689; PIDN:AAL43996.1;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu3180
A/Map position: linear chromosome

Query Match 52.0%; Score 51; DB 2; Length 277;
Best Local Similarity 58.8%; Pred. No. 1.5;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

CY 4 DNAVLEGS LVKVTGANG 20
Db 25 DGAIRGSLTAVVGANG 41

RESULT 2
D98335
ABC transporter, ATP-binding protein nmb0588 [imported] - Agrobacterium tumefaciens (str
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: D98335
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: D98335
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-277 <KUR>
A/Cross-references: UNIPROT:Q8UB37; UNIPARC:UPI00000D20E7; GB:AE007870; PIDN:AAK90206.1;
C/Genetics:
A/Gene: AGR_L_3258
A/Map position: linear chromosome

A;Residues: 1-111 <STO>
A;Cross-references: UNIPROT:Q8X9P2; UNIPARC:UPI00000D07CC; GB:AE005174; NID:g12513991; H
A;Experimental source: strain O157:H7, substrain EDL933
A;Accession: E85659
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-111 <ST2>
A;Cross-references: UNIPARC:UPI00000D07CC; GB:AE005174; NID:g12514511; PIDN:AAG55737.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1183; Z1622

Query Match 48.0%; Score 47; DB 2; Length 111;
Best Local Similarity 41.2%; Pred. No. 2.5;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 PDNAVLEGLVKVTGAN 19
|||::|:|
Db 28 PDNAIIQKKAISVGRN 44

RESULT 8

D90218
1SU ribosomal protein L6AB (rpl6AB) [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: D90218
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Senen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139

A;Accession: D90218
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-181 <KUR>
A;Cross-references: UNIPROT:Q9UX91; UNIPARC:UPI000013416E; GB:AE006641; NID:g13813871; F
C;Genetics:
A;Gene: rpl6AB
C;Superfamily: ribosomal protein L6/L9

Query Match 48.0%; Score 47; DB 2; Length 181;
Best Local Similarity 47.6%; Pred. No. 4.2;
Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

OY 2 IPDNAV--LEGLVKVTGAN 20
|||::|:|
Db 12 IPKVVVVDLKGSIKIKGPKG 32

RESULT 9

H70326
hypothetical protein aq_294 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: H70326
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: H70326
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-371 <AQF>
A;Cross-references: UNIPROT:O66644; UNIPARC:UPI00000562FF; GB:AE000682; NID:g2982979; PI
A;Experimental source: strain VFS
C;Genetics:
A;Gene: aq_294

Query Match 46.9%; Score 46; DB 2; Length 371;
Best Local Similarity 57.1%; Pred. No. 13;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 IPDNAVLEGLVKV 15
||:|::|:|
Db 321 IPEKALFGRNVKI 334

RESULT 10

T06029
hypothetical protein T28119.100 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06029
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15484
A;Accession: T06029
A;Molecule type: DNA
A;Residues: 1-532 <BEV>
A;Cross-references: UNIPROT:Q9T069; UNIPARC:UPI00000ABF4D; EMBL:AL035709; GSPDB:GN00062;
A;Experimental source: cultivar Columbia; BAC clone T28119
C;Genetics:
A;Gene: ATSP:T28119.100
A;Map position: 4

Query Match 46.9%; Score 46; DB 2; Length 532;
Best Local Similarity 52.9%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 PDNAVLEGLVKVTGAN 19
|||::|:|
Db 209 PKNEVLEGSVIRKESLN 225

RESULT 11

T32934
hypothetical protein W03D8.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T32934
R;Jones, K.; Graves, T.; Ozersky, P.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid W03D8.
A;Reference number: Z21249
A;Accession: T32934
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1647 <JON>
A;Cross-references: UNIPARC:UPI0000164013; EMBL:AF043702; PIDN:AAB97583.1; GSPDB:GN00019
A;Experimental source: strain Bristol N2; clone W03D8
C;Genetics:
A;Gene: CESP:W03D8.6
A;Map position: 1
A;Introns: 47/1; 82/3; 115/1; 230/1; 319/2; 406/3; 466/1; 513/1; 554/2; 586/3; 653/1; 70
1589/2

Query Match 46.9%; Score 46; DB 2; Length 1647;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 IPDNAVLEGLV 13
|||::|:|
Db 808 IGDNAVLEGSVV 819

RESULT 12

A32687
lysyl endopeptidase (EC 3.4.21.50) precursor - Achromobacter lyticus (strain M497-1)
N;Alternate names: Achromobacter proteinase I
C;Species: Achromobacter lyticus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: A32687; A32960
R;Ohara, T.; Makino, K.; Shinagawa, H.; Nakata, A.; Norioka, S.; Sakiyama, F.

J. Biol. Chem. 264, 20625-20631, 1989
A>Title: Cloning, nucleotide sequence, and expression of Achromobacter protease I gene.
A/Reference number: A32687; MUID:90062204; PMID:2684982
A/Accession: A32687
A/Molecule type: DNA
A/Residues: 1-653 <OHA>
A/Cross-references: UNIPROT:P15636; UNIPARC:UPI0000125C8B; GB:J05128; NID:g141803; PIDN:R.Tsunasawa, S.; Masaki, T.; Hirose, M.; Soejima, M.; Sakiyama, F.
J. Biol. Chem. 264, 3832-3839, 1989
A>Title: The primary structure and structural characteristics of Achromobacter lyticus F
A/Reference number: A32960; MUID:89139439; PMID:2492988
A/Accession: A32960
A/Molecule type: protein
A/Residues: 206-473 <TSU>
A/Cross-references: UNIPARC:UPI000011061E
C/Comment: This protein specifically hydrolyzes lysyl peptide bonds including the lys-Px
C/Keywords: hydrolase; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-205/Domain: propeptide #status predicted <PRO>
F:206-473/Product: achromobacter proteinase I #status experimental <MAT>
F:211-421,217-285,241-263/Disulfide bonds: #status experimental
F:262,318,399/Active site: His, Asp, Ser #status predicted

Query Match 46.4%; Score 45.5; DB 2; Length 653;
Best Local Similarity 55.0%; Pred. No. 29;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 1 AIPDNAVLEGSIVKVTGANG 20
| | | | | : | : | : | : |
Db 569 AIPDNATVE-SPIVSGRTG 587

RESULT 13
T33939
Probable alcohol dehydrogenase (EC 1.1.1.1) [similarity] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T33939
R/Becker, M.; Graves, T.; Wilson, C.
submitted to the EMBL Data Library, February 1999
A/Description: The sequence of C. elegans cosmid D2063.
A/Reference number: 221442
A/Accession: T33939
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-408 <BEC>
A/Cross-references: UNIPROT:Q9UAT1; UNIPARC:UPI00000762B1; EMBL:AF125951; P1DN:AAD14690.
A/Experimental source: strain Bristol N2; clone D2063
C/Genetics:
A/Gene: CESP:D2063.1
A/Map position: 5
A/Introns: 294/3; 359/3; 373/3
C/Superfamily: alcohol dehydrogenase, long-chain alcohol dehydrogenase homology
C/Keywords: NAD; oxidoreductase; zinc
F:47,70,173/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted

Query Match 45.9%; Score 45; DB 2; Length 408;
Best Local Similarity 45.0%; Pred. No. 21;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AIPDNAVLEGSIVKVTGANG 20
| : : | : | | | | |
Db 181 ALKESEVKSQIVAVTGAGG 200

RESULT 14
B46230
RNA-binding protein homolog Cpo (clone 61.2) - fruit fly (Drosophila melanogaster)
N/Alternate names: couch potato protein 61.2; cpo61.2 protein
C/Species: Drosophila melanogaster
C/Date: 21-Sep-1993 #sequence_revision 30-Apr-1999 #text_change 15-Mar-2004
C/Accession: B46230; S24678
R/Bellen, H.J.; Kooyer, S.; D'Evelyn, D.; Pearlman, J.

Genes Dev. 6, 2125-2136, 1992
A>Title: The Drosophila couch potato protein is expressed in nuclei of peripheral neurons
A/Reference number: A46230; MUID:93051332; PMID:1427076
A/Accession: B46230
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-536 <BEL>
A/Cross-references: UNIPARC:UPI000017BECA; GB:Z14312
A/Note: sequence extracted from NCBI backbone (NCBIN:117907, NCBI:P:117908)
R/Bellen, H.J.
submitted to the EMBL Data Library, August 1992
A/Accession: S24678
A/Molecule type: DNA
A/Residues: 1-177,'T',179-311,'T',313-319,'G',321-536 <BEF>
A/Cross-references: UNIPARC:UPI000017BECB; EMBL:Z14312
C/Genetics:
A/Gene: FlyBase:cpo
A/Start codon: CTC
C/Keywords: alternative splicing, RNA binding
F:453-521/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 45.9%; Score 45; DB 2; Length 536;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 EGSLVKTGANG 20
| | | | | : | | | | |
Db 479 EGSLKVTSKNG 490

RESULT 15
S24761
Cpo 61.1 protein - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Mar-2004
C/Accession: S24761
R/Bellen, H.J.
submitted to the EMBL Data Library, August 1992
A/Reference number: S2477
A/Accession: S24761
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-615 <BEL>
A/Cross-references: UNIPARC:UPI000017BE84; EMBL:Z14974
C/Genetics:
A/Gene: FlyBase:cpo
A/Cross-references: FlyBase:FBgn0000363
A/Start codon: CTC
A/Introns: 450/1; 475/1; 493/3; 514/3; 562/1
F:452-520/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 45.9%; Score 45; DB 2; Length 615;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 EGSLVKTGANG 20
| | | | | : | | | | |
Db 478 EGSLKVTSKNG 489

Search completed: March 11, 2006, 05:09:43
Job time : 20.4127 secs

DB 244 AIPESLVLEGHLEKAFSRNG 263

RESULT 4

```
Q65DJ1_BACLD PRELIMINARY; PRT; 396 AA.
ID Q65DJ1; Q62P10;
AC Q65DJ1; Q62P10;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE ywbD (SAM (And some other nucleotide) binding motif, PUA domain).
GN Name=ywbD; OrderedLocuNames=BL03857, BL104060;
OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15383718; DOI=10.1159/000079829;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The complete genome sequence of Bacillus licheniformis DSM13, an
RT organism with great industrial potential."
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., Lopez de Leon A., Xiang H., Gustl V., Clausen I.G.,
RA Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA Larsen T.S., Sorokin A., Belotkin A., Lapidus A., Galleron N.,
RA Ehrlich S.D., Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
DR EMBL; AB017333; AAU42873.1; -; Genomic_DNA.
DR EMBL; CP000002; AAU25501.1; -; Genomic_DNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR InterPro; IPR002478; PUA.
DR InterPro; IPR000051; SAM_bd.
DR SMART; SM00359; PUA; 1.
DR PROSITE; PS50890; PUA; 1.
KW Complete proteome.
SQ SEQUENCE 396 AA; 44897 MW; C1FFD1AC5EEFE2AE CRC64;
```

Query Match 52.0%; Score 51; DB 2; Length 396;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 4 DNAVLEGSIVKVTGANG 20
Db 30 ENGIRKESLIRITDENG 46

RESULT 5

```
Q9MM39_9SPER PRELIMINARY; PRT; 412 AA.
ID Q9MM39_9SPER PRELIMINARY; PRT; 412 AA.
AC Q9MM39;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE F1 ATPase alpha subunit (Fragment).
OS Name=atpA;
OS Gnetum ula.
OS Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales;
OX NCBI_TaxID=3383;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20226063; PubMed=10760278; DOI=10.1073/pnas.97.8.4092;
```

RA Bowe L.M., Coat G., dePamphilis C.W.;
RT "Phylogeny of seed plants based on all three genomic compartments:
RT extant gymnosperms are monophyletic and Gnetales' closest relatives
RT are conifers.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:4092-4097(2000).

```
CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
DR EMBL; AF209109; AAF64667.1; -; Genomic_DNA.
DR HSSP; P19483; 1H8E.
DR SMR; Q9MM39; 1-399.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0045261; C:proton-transporting ATP synthase complex, C. . .; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity. . .; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. . .; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR GO; GO:0006811; P:proton transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR005294; ATP_synthF1_alph.
DR InterPro; IPR004100; ATPase_a/b_N.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR Pfam; PF00006; ATP-synt_ab; 1.
DR Pfam; PF02874; ATP-synt_ab_N; 1.
DR TIGRFAMs; TIGR00962; atpA; 1.
DR PROSITE; PS00152; ATPASE ALPHA BETA; 1.
KW ATP synthesis; ATP-binding; CF(1); Hydrogen ion transport; Hydrolase;
KW ion transport; Mitochondrion; Nucleotide-binding; Transport.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 412 AA; 43993 MW; 824873B6ED21327B CRC64;
```

Query Match 51.0%; Score 50; DB 2; Length 412;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 4 DNAVLEGSIVKVTGA 18
Db 45 DTAIRKESIVKRTGS 59

RESULT 6

```
Q9T7A3_GNEGN PRELIMINARY; PRT; 418 AA.
ID Q9T7A3_GNEGN PRELIMINARY; PRT; 418 AA.
AC Q9T7A3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE ATPase alpha subunit (Fragment).
OS Name=atp1;
OS Gnetum gnemon (Bago).
OS Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales;
OX NCBI_TaxID=3382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20052162; PubMed=10586879; DOI=10.1038/46536;
RA Qiu Y.L., Lee J., Bernasconi-Quadrone F., Soltis D.B., Soltis P.S.,
RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
RT "The earliest angiosperms: evidence from mitochondrial, plastid and
RT nuclear genomes."
RL Nature 402:404-407(1999).
CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
DR EMBL; AF197617; AAF16949.1; -; Genomic_DNA.
DR HSSP; P19483; 1E79.
DR SMR; Q9T7A3; 1-418.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0045261; C:proton-transporting ATP synthase complex, C. . .; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity. . .; IEA.
```

DR	GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. . .; IEA.
DR	GO; GO:0016787; F:hydrolase activity; IEA.
DR	GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR	GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR	GO; GO:0006811; P:ion transport; IEA.
DR	GO; GO:0015992; P:proton transport; IEA.
DR	InterPro; IPR005294; ATP synthet1_alph.
DR	InterPro; IPR000793; ATPase_a/b_C.
DR	InterPro; IPR004100; ATPase_a/b_N.
DR	InterPro; IPR000194; ATPase_a/b_centre.
DR	Pfam; PF00006; ATP-synt_ab; 1.
DR	Pfam; PF00306; ATP-synt_ab_C; 1.
DR	Pfam; PF02874; ATP-synt_ab_N; 1.
DR	TIGRfams; TIGR00962; atpA; 1.
DR	PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW	ATP synthesis; ATP-binding; C(1); Hydrogen ion transport; Hydrolase;
KM	ion transport; Mitochondrion; Nucleotide-binding; Transport.
FT	NON_TER 1
FT	NON_TER 1
FT	NON_TER 1
SEQ	SEQUENCE 418 AA; 44753 MW; 0FE555A3B5722901 CRC64;
QY	Query Match 51.0%; Score 50; DB 2; Length 418;
DB	Best Local Similarity 60.0%; Pred. No. 48;
DB	Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
DB	4 DNAVLEGS LKVTGA 18
DB	44 DTAIEGSIVKRTGS 58

```

RESULT 7
Q69IP6 ORYSA
ID Q69IP6_ORYSA PRELIMINARY; PRT; 451 AA.
AC Q69IP6;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Putative aspartic proteinase nepenthesin I.
GN Name=OSJNBa0039D04.13;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RL clone:OSJNBa0039D04.";
RT Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006753; BAD32124.1; -, Genomic_DNA.
DR Gramene; Q69IP6; -.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00026; Asp_1.
DR PRINTS; PRO0792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN 1.
DR SEQ SEQUENCE 451 AA; 46440 MW; E7F24FED78353C5E CRC64;

Query Match 51.0%; Score 50; DB 2; Length 451;
Best Local Similarity 62.5%; Pred. No. 51;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 4 DNAYLEGSLVKVTGAN 19
| : : | | | | | | |
DB 243 DSPILFGSLAKVTGCN 258

RESULT 8
Q4RSV8 TETNG
ID Q4RSV8_TETNG PRELIMINARY; PRT; 687 AA.

```

AC QARSV8; 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Chromosome 12 SCAFL14999, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00029526001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OX Tetraodontidae; Tetraodontidae; Tetraodon.
RN NCB1_TaxID=99883;
RN [1]
RP NTCTEOTIDE SEQUENCE.
RA Jallion O., Auray J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Iutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Archonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Couranceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NTCTEOTIDE SEQUENCE.
RG GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014999; CAG08524.1; -; Genomic_DNA.
FT NON_TER 687 687
SQ SEQUENCE 687 AA; 72523 MW; 3A4531EE8D7AB537 CRC64;

Query Match 51.0%; Score 50; DB 2; Length 687;
Best Local Similarity 40.0%; Pred. No. 78;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 AIPDNAVLGSLVKVTGANG 20
::||| : |||: ||| |
Db 356 SVPPNVTAGSVIPIVTGGRG 375

```

RESULT 9
O6BIS4_DEBHA
ID Q6BIS4_DEBHA PRELIMINARY; PRT; 1048 AA.
AC Q6BIS4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to CA1657|IPF16022 Candida albicans IPF16022 unknown
DE function.
GN OrderedLocusNames=DEHA0G08679g;
OS Debaryomyces hanseni (Yeast) (Torulasporea hanseni).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 36239 / CBS 767;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barney S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boistrame A., Boyer J., Cattolico L., Confanioleri F., de Darvar A.,

```



```
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hanttraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Kozsul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.-M., Nikolski M., Oztas S., Ozler-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.,
RT "Genome evolution in yeasts.",
RL Nature 430:35-44(2004).
DR EMBL; CR382139; CAG90360.1; -; Genomic_DNA.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Complete proteome; Repeat; WD repeat.
SQ SEQUENCE 1048 AA; 117153 MW; 7BAEB602ACF85373 CRC64;

Query Match          51.0%; Score 50; DB 2; Length 1048;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 PDNAVLEGLVKVTGANG 20
DB 26 PDNFVIEGNLMAYTAGG 43

RESULT 10
O8G456 BIFLO PRELIMINARY; PRT; 3172 AA.
AC O8G456;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fab.
GN Name=fas; OrderedLocustNames=BL1537;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Scheil M.A., Karimantzou M., Snel B., Vilanova D., Berger B.,
RA Peest G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014295; AAN25329.1; -; Genomic_DNA.
DR HSSP; P25715; 1MLA.
DR GO; GO:0005835; C:fatty acid synthase complex; IEA.
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR004136; 2nprop_dioxygen.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR003965; Fatty_acid_synth.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR002539; Maoc_dehydratase.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00109; ketoacyl_synt; 1.
DR Pfam; PF02801; ketoacyl_synt_C; 1.
DR Pfam; PF01575; Maoc_dehydratase; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Complete proteome.
```

```
SQ SEQUENCE 3172 AA; 338131 MW; 351478B6BF20CFDE CRC64;

Query Match          51.0%; Score 50; DB 2; Length 3172;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGLVKVTGANG 20
DB 2026 AVVDSALDAFAEKVTGSNG 2045

RESULT 11
O9HYS1 PSEAE PRELIMINARY; PRT; 592 AA.
AC O9HYS1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable short-chain dehydrogenase.
GN OrderedLocustNames=PA3324;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004754; AAG06712.1; -; Genomic_DNA.
DR PIR; B83231; B83231.
DR HSSP; O724W1; 1PR9.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR InterPro; IPR000639; Epox_hydrolase.
DR InterPro; IPR000379; Ser_esters.
DR Pfam; PF00561; Abhydrolase_1; 1.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00412; EPOXH_YDR_LAS_E.
DR PRINTS; PR00081; GDHRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 592 AA; 65770 MW; 7EA0D0996D4D0157 CRC64;

Query Match          50.0%; Score 49; DB 2; Length 592;
Best Local Similarity 55.6%; Pred. No. 96;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 PDNAVLEGLVKVTGANG 20
DB 318 PDASSMNGKLVVVTGAGG 335

RESULT 12
Q4ZUUS PSESY PRELIMINARY; PRT; 732 AA.
AC Q4ZUUS;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Outer membrane autotransporter barrel precursor.
```

GN ORFNames=Psy_2034;
OS Pseudomonas syringae pv. syringae B728a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205918;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG DOE Joint Genome Institute;
RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
RA Nolan M., Goldsman E., Thiel J., Malfatti S., Lapidus A., Detter J.C.,
RA Land M., Richardson P.M., Kyrpides N.C., Ivanova N.;
RT "Comparison of two complete genome sequences of Pseudomonas syringae
pv. syringae B728a and pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Loper J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Reil H., Fell W.S., Lindow S.E.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000075; AAAY37077.1; -; Genomic_DNA.
DR InterPro; IPR005546; Auto_transptbeta.
DR InterPro; IPR006315; Auto transporter.
DR InterPro; IPR004899; Pertactin.
DR InterPro; IPR003991; Pertactin_C.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF03212; Pertactin; 1.
DR PRINTS; PR01484; PRTACTNFAMILY.
DR TIGRFAMs; TIGR01414; auto trans_bar1; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 732 AA; 75763 MW; 4C355FC1A478A628 CRC64;

Query Match 50.0%; Score 49; DB 2; Length 732;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 NAVLEGSIVKVTGANG 20
Db 148 NLTEGTSVTATGANG 163

RESULT 13
DNL1_PYRFU
ID DNL1_PYRFU STANDARD; PRT; 561 AA.
AC P56709;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Thermosable DNA ligase (EC 6.5.1.1) (Polydeoxyribonucleotide synthase
[ATP]) (Pfu DNA ligase).
GN Name=lig; Ordered locus Names=PF1635;
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF N-TERMINUS.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Mathur E.J., Marsh E.J., Schoettlin W.E.;
RT "Purified thermostable Pyrococcus furiosus DNA ligase.";
RL Patent number US5700672, 23-DEC-1997.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: This protein seals, during DNA replication, DNA
CC recombination and DNA repair, nicks in double-stranded DNA.
CC -!- CATALYTIC ACTIVITY: ATP + (deoxyribonucleotide) (n) +
CC (deoxyribonucleotide) (m) = AMP + diphosphate +
CC (deoxyribonucleotide) (n+m).
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
CC Temperature dependence:
CC Optimum temperature is about 70 degrees Celsius. Active from 4
CC to 100 degrees Celsius. Thermostable;
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE010263; AAL81759.1; -; Genomic_DNA.
DR HAMAP; MF_00407; -; 1.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR012309; DNA_ligase_A_C.
DR InterPro; IPR012310; DNA_ligase_A_M.
DR InterPro; IPR012308; DNA_ligase_A_N.
DR InterPro; IPR012340; OB_NA_bd_sub.
DR PANTHER; PTHR10459; DNA_ligase; 1.
DR Pfam; PF04679; DNA_ligase_A_C; 1.
DR Pfam; PF01068; DNA_ligase_A_M; 1.
DR Pfam; PF04675; DNA_ligase_A_N; 1.
DR TIGRFAMs; TIGR00574; dnl1; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR PROSITE; PS00333; DNA_LIGASE_A2; FALSE_NEG.
DR PROSITE; PS50160; DNA_LIGASE_A3; 1.
KW ATP-binding; Cell cycle; Cell division; Complete proteome;
KW Direct protein sequencing; DNA damage; DNA recombination; DNA repair;
KW DNA replication; ligase; Nucleotide-binding.
FT ACT_SITE 249 249 N6-AMP-lysine intermediate (By
FT similarity).
SQ SEQUENCE 561 AA; 63773 MW; 534158525B9D24B2 CRC64;

Query Match 49.5%; Score 48.5; DB 1; Length 561;
Best Local Similarity 52.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 2 IPDNAVLEGSIVKVTGANG 20
Db 290 IPEKATVEGELVAI-GENG 307

RESULT 14
Q6AHE3_LEIXX
ID Q6AHE3_LEIXX PRELIMINARY; PRT; 340 AA.
AC Q6AHE3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Alcohol dehydrogenase.
GN Name=adhT; Ordered locus Names=lx01190;
OS Leifsonia xyl1 (subsp. xyl1).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F., Jr., Carer H., Coutinho L.L., El-Dorry H.A.,
RA Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,

Job time : 122.095 secs

RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
 RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
 RA Teal S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.;
 RT "The genome sequence of the Gram-positive sugarcane pathogen *Leifsonia*
 RT *xyli* subsp. *xyli*."
 RL Mol. Plant Microbe Interact. 17:827-836(2004).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 DR EMBL: AE016822; AAT88202.1; -; Genomic DNA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR InterPro: IPR002085; ADH_SF_Zn.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 KW Complete proteome.
 SQ SEQUENCE 340 AA; 35189 MW; D8CD56AB6F529543 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 340;
 Best Local Similarity 60.0%; Pred. No. 79;
 Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AIPDNAVLEGS LVKVTGANG 20
 Db 151 ADPDRVRAEGSSVLTGAAG 170

RESULT 15
 Q97AQ8 THEVO PRELIMINARY; PRT; 404 AA.
 ID Q97AQ8 THEVO
 AC Q97AQ8;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Carboxypeptidase.
 GN OrderedLocuNames=TV0751; ORFNames=TVG0758219;
 OS Thermoplasma volcanum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasmata.
 OC NCBI_TaxID=50339;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=GSS1 / DSM 4299 / JCM 9571;
 RX MEDLINE=20570466; PubMed=11121031; DOI=10.1073/pnas.97.26.14257;
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
 RA Kawashima T., Yamamoto Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
 RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
 RT "Archaeal adaptation to higher temperatures revealed by genomic
 RT sequence of *Thermoplasma volcanum*."
 RT Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
 RL EMBL: BA000011; BAB59893.1; -; Genomic DNA.
 DR MEROPS; M20.008; -;
 DR GO: GO:0004180; F:carboxypeptidase activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:0008237; F:metallopeptidase activity; IEA.
 DR GO: GO:0046983; F:protein dimerization activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR011650; M20_dimer.
 DR InterPro: IPR002933; Peptidase_M20.
 DR InterPro: IPR010168; Pept_M20D_amidh.
 DR Pfam: PF07687; M20_dimer; 1.
 DR Pfam: PF01546; Peptidase_M20; 1.
 DR TIGRFAMs: TIGR01891; amidohydrolases; 1.
 KW Carboxypeptidase; Complete proteome.
 SQ SEQUENCE 404 AA; 44080 MW; 4D6BBE73321C7327 CRC64;

Query Match 49.0%; Score 48; DB 2; Length 404;
 Best Local Similarity 81.8%; Pred. No. 94;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 IIPDNAVLEGS L 12
 Db 261 IIPDNAVLEGS L 271

This Page Blank (uspic,

```
RESULT 4
AD056196
ID   AD056196 standard; protein; 343 AA.
XX
XX
AC   AD056196;
XX
DT   12-AUG-2004 (first entry)
XX
XX   Sporobolomyces salmonicolor ketoreductase, aldehyde reductase II.
DE
XX   (R)-hydroxy ester; alpha-keto ester; ketoreductase; enzyme;
KM   aldehyde reductase II.
XX
OS   Sporidiobolus salmonicolor.
XX
PN   US2004101937-A1.
XX
PD   27-MAY-2004.
XX
PF   08-JUL-2003; 2003US-00616320.
XX
PR   10-JUL-2002; 2002US-0394761P.
XX
PA   (MOOR/) MOORE J C.
PA   (STUR/) STURR M G.
PA   (MCLA/) MCLAUGHLIN K.
PA   (KIMJ/) KIM J.
XX
PI   Moore JC, Sturr MG, McLaughlin K, Kim J;
XX
DR   WPI; 2004-431257/40.
XX
PT   Preparing (R)-hydroxy ester involves adding alpha-keto ester to mixture
PT   comprising ketoreductase enzyme and non-ketoreductase enzyme components.
XX
PS   Disclosure; SEQ ID NO 3; 10pp; English.
XX
XX   The invention relates to a method of preparing high yields of an (R)-
CC   hydroxy ester by reducing an alpha-keto ester using a ketoreductase
CC   enzyme. The present sequence represents Sporobolomyces salmonicolor
CC   ketoreductase, aldehyde reductase II used in the method of the invention.
XX
SQ   Sequence 343 AA;

Query Match          100.0%; Score 107; DB 8; Length 343;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   1 MAKIDNAVLPEGSLVLTGTANG 22
    |||||
Db   1 MAKIDNAVLPEGSLVLTGTANG 22

RESULT 5
AD056197
ID   AD056197 standard; peptide; 21 AA.
XX
XX
AC   AD056197;
XX
DT   12-AUG-2004 (first entry)
XX
XX   Sporobolomyces salmonicolor aldehyde reductase II, N-terminal peptide #3.
DE
XX   (R)-hydroxy ester; alpha-keto ester; ketoreductase; enzyme;
KM   aldehyde reductase II.
XX
OS   Sporidiobolus salmonicolor.
XX
PN   US2004101937-A1.
XX
PD   27-MAY-2004.
XX
```

```
PF   08-JUL-2003; 2003US-00616320.
XX
XX
PR   10-JUL-2002; 2002US-0394761P.
XX
XX
PA   (MOOR/) MOORE J C.
PA   (STUR/) STURR M G.
PA   (MCLA/) MCLAUGHLIN K.
PA   (KIMJ/) KIM J.
XX
PI   Moore JC, Sturr MG, McLaughlin K, Kim J;
XX
DR   WPI; 2004-431257/40.
XX
PT   Preparing (R)-hydroxy ester involves adding alpha-keto ester to mixture
PT   comprising ketoreductase enzyme and non-ketoreductase enzyme components.
XX
PS   Claim 1; SEQ ID NO 4; 10pp; English.
XX
XX   The invention relates to a method of preparing high yields of an (R)-
CC   hydroxy ester by reducing an alpha-keto ester using a ketoreductase
CC   enzyme. The present sequence represents a Sporobolomyces salmonicolor
CC   ketoreductase, aldehyde reductase II, N-terminal peptide used in the
-CC   method of the invention.
XX
SQ   Sequence 21 AA;

Query Match          64.0%; Score 68.5; DB 8; Length 21;
Best Local Similarity 81.8%; Pred. No. 0.0011;
Matches 18; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY   1 MAKIDNAVLPEGSLVLTGTANG 22
    |||||
Db   1 MAIPDNAVL-EGSLVKTGTANG 21

RESULT 6
AD056194
ID   AD056194 standard; peptide; 20 AA.
XX
XX
AC   AD056194;
XX
DT   12-AUG-2004 (first entry)
XX
XX   Sporobolomyces salmonicolor aldehyde reductase II, N-terminal peptide #1.
DE
XX   (R)-hydroxy ester; alpha-keto ester; ketoreductase; enzyme;
KM   aldehyde reductase II.
XX
OS   Sporidiobolus salmonicolor.
XX
PN   US2004101937-A1.
XX
PD   27-MAY-2004.
XX
PF   08-JUL-2003; 2003US-00616320.
XX
PR   10-JUL-2002; 2002US-0394761P.
XX
PA   (MOOR/) MOORE J C.
PA   (STUR/) STURR M G.
PA   (MCLA/) MCLAUGHLIN K.
PA   (KIMJ/) KIM J.
XX
PI   Moore JC, Sturr MG, McLaughlin K, Kim J;
XX
DR   WPI; 2004-431257/40.
XX
PT   Preparing (R)-hydroxy ester involves adding alpha-keto ester to mixture
PT   comprising ketoreductase enzyme and non-ketoreductase enzyme components.
XX
PS   Claim 1; SEQ ID NO 1; 10pp; English.
XX
XX   The invention relates to a method of preparing high yields of an (R)-
```

CC hydroxy ester by reducing an alpha-keto ester using a ketoreductase
CC enzyme. The present sequence represents a Sporobolomyces salmonicolor
CC ketoreductase, aldehyde reductase II, N-terminal peptide used in the
CC method of the invention.
XX

SQL Sequence 20 AA;

Query Match 61.2%; Score 65.5; DB 8; Length 20;
Best Local Similarity 88.9%; Pred. No. 0.0032;
Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 5 DNAVLPESLVLTGANG 22
|||||
DB 4 DNAVLP-EGSLVKVTGANG 20

RESULT 7
ADS25047
ID ADS25047 standard; protein; 323 AA.

AC ADS25047;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #14080.

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

DR WPI; 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 14080; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQL Sequence 323 AA;

Query Match 50.5%; Score 54; DB 8; Length 323;
Best Local Similarity 59.1%; Pred. No. 6.2;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 MAKIDNAVLPESLVLTGANG 22
|||:|||||
DB 137 MALERNGVTPOGGEVLVTGANG 158

RESULT 8
AAV38793
ID AAV38793 standard; protein; 461 AA.

AC AAV38793;

DT 08-OCT-1999 (first entry)

DE Neisseria gonorrhoeae antigenic protein encoded by ORF140.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.

OS Neisseria gonorrhoeae.

PN WO9924578-A2.

PD 20-MAY-1999.

PF 09-OCT-1998; 98WO-IB001665.

PR 06-NOV-1997; 97GB-00023516.

PR 14-NOV-1997; 97GB-00024190.

PR 18-NOV-1997; 97GB-00024386.

PR 27-NOV-1997; 97GB-00025158.

PR 10-DEC-1997; 97GB-00026147.

PR 14-JAN-1998; 98GB-00000759.

PR 01-SEP-1998; 98GB-00019016.

PA (CHIR-) CHIRON SPA.

PI Massignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;

DR WPI; 1999-327407/27.

PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection.

PS Claim 4; Page 336; 524pp; English.

XX Amino acid sequences AAV38499-Y38944 represent Neisseria meningitidis and
CC N. gonorrhoeae antigenic proteins. They are encoded by open reading
CC frames (ORFs) AA21972-Z12358. The antigenic proteins, their fragments,
CC their nucleic acids and antibodies are used for diagnosis, prevention (as
CC vaccines) or treatment of Neisseria infections, such as meningitis,
CC septicemia and gonorrhea. Both organisms are closely related. Fragments
CC of the nucleic acids are useful as hybridisation probes and antisense
XX reagents
SQL Sequence 461 AA;

Query Match 48.6%; Score 52; DB 2; Length 461;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPESGLVLTGANG 22
:|:|:|:|:|:|
Db 312 VDGALAPACSVILITGAGG 330

RESULT 9
AAV38794
ID AAV38794 standard; protein; 461 AA.
XX
AC AAV38794;
XX
DT 08-OCT-1999 (first entry)
XX
DE N. gonorrhoeae antigenic protein encoded by variant ORF140.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.
XX
OS Neisseria gonorrhoeae.
XX
PN WO924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-1B001665.
XX
PR 06-NOV-1997; 97GB-00023516.
PR 14-NOV-1997; 97GB-00024190.
PR 18-NOV-1997; 97GB-00024386.
PR 27-NOV-1997; 97GB-00025158.
PR 10-DEC-1997; 97GB-00026147.
PR 14-JAN-1998; 98GB-00000759.
PR 01-SEP-1998; 98GB-00019016.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX
DR WPI; 1999-327407/27.
DR N-PSDB; AA212227.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection.
XX
PS Claim 4; Page 336; 524pp; English.
XX
CC Amino acid sequences AAV38499-Y38944 represent Neisseria meningitidis and
CC N. gonorrhoeae antigenic proteins. They are encoded by open reading
CC frames (ORFs) AA211972-212358. The antigenic proteins, their fragments,
CC their nucleic acids and antibodies are used for diagnosis, prevention (as
CC vaccines) or treatment of Neisseria infections, such as meningitis,
CC septicemia and gonorrhea. Both organisms are closely related. Fragments
CC of the nucleic acids are useful as hybridisation probes and antisense
CC reagents
XX
SO Sequence 461 AA;

Query Match 48.6%; Score 52; DB 2; Length 461;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPESGLVLTGANG 22
:|:|:|:|:|:|
Db 312 VDGALAPACSVILITGAGG 330

RESULT 10
AEB49406

ID AEB49406 standard; protein; 461 AA.
XX
AC AEB49406;
XX
DT 22-SEP-2005 (first entry)
XX
DE N. gonorrhoeae protein SEQ ID NO 592.
XX
KW antibacterial; immunosuppressive; antiinflammatory; vaccine;
KW neisseria gonorrhoeae infection; gynecological; infection;
KW neisseria meningitidis infection; meningitis; neuroprotective;
KW inflammation; neurological disease; sepsis; diagnosis.
XX
OS Neisseria gonorrhoeae.
XX
PN US6914131-B1.
XX
PD 05-JUL-2005.
XX
PF 30-APR-1999; 99US-00303518.
XX
PR 09-OCT-1998; 98WO-1B001665.
XX
PA (CHIR) CHIRON SRL.
XX
PI Scarlato V, Masignani V, Rappuoli R, Pizza M, Grandi G;
XX
DR WPI; 2005-464877/47.
DR N-PSDB; AEB49405.
XX
PT New Neisserial nucleic acids useful for diagnosing and/or treating
PT bacterial infections, in particular meningitis and septicemia caused by
PT Neisseria meningitidis and Neisseria gonorrhea.
XX
PS Example 70; SEQ ID NO 592; 613pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule. The methods
CC and compositions of the present invention are useful for diagnosing
CC and/or treating Neisserial bacterial infections, in particular meningitis
CC and septicemia caused by Neisseria meningitidis and Neisseria gonorrhea.
CC The present sequence represents the amino acid sequence of a N.
CC gonorrhoeae protein. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6914131B1.
XX
SO Sequence 461 AA;

Query Match 48.6%; Score 52; DB 9; Length 461;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPESGLVLTGANG 22
:|:|:|:|:|:|
Db 312 VDGALAPACSVILITGAGG 330

RESULT 11
AEB49404
ID AEB49404 standard; protein; 461 AA.
XX
AC AEB49404;
XX
DT 22-SEP-2005 (first entry)
XX
DE N. gonorrhoeae protein SEQ ID NO 590.
XX
KW antibacterial; immunosuppressive; antiinflammatory; vaccine;
KW neisseria gonorrhoeae infection; gynecological; infection;
KW neisseria meningitidis infection; meningitis; neuroprotective;
KW inflammation; neurological disease; sepsis; diagnosis.
XX
OS Neisseria gonorrhoeae.
XX

PN US6914131-B1.
XX
PD 05-JUL-2005.
XX
PF 30-APR-1999; 99US-00303518.
XX
PR 09-OCT-1998; 98WO-IB001665.
XX
PA (CHIR) CHIRON SRL.
XX
PI Scarlato V, Maignani V, Rappuoli R, Pizza M, Grandi G;
XX
DR WPI; 2005-464877/47.
XX
PT New Neisserial nucleic acids useful for diagnosing and/or treating
PT bacterial infections, in particular meningitis and septicemia caused by
PT Neisseria meningitidis and Neisseria gonorrhea.
XX
PS Example 70; SEQ ID NO 590; 613pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule. The methods
CC and compositions of the present invention are useful for diagnosing
CC and/or treating Neisserial bacterial infections, in particular meningitis
CC and septicemia caused by Neisseria meningitidis and Neisseria gonorrhea.
CC The present sequence represents the amino acid sequence of a N.
CC gonorrhoeae protein. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?docID=6914131B1.
XX
SQ Sequence 461 AA;

Query Match 48.6%; Score 52; DB 9; Length 461;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPEGSLVLTGANG 22
:| |:| |:| |:| |:| |
Db 312 VDGALAPACSVLITGAGG 330

RESULT 12
ABP77208
ID ABP77208 standard; protein; 485 AA.
XX
AC ABP77208;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 946.
XX
KW Antibacterial; infection; vaccine; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Maignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
DR N-PSDB; ABZ38178.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
XX
PS Disclosure; Page 257; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 485 AA;

Query Match 48.6%; Score 52; DB 6; Length 485;
Best Local Similarity 47.4%; Pred. No. 21;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPEGSLVLTGANG 22
:| |:| |:| |:| |:| |
Db 336 VDGALAPACSVLITGAGG 354

RESULT 13
AAY38791
ID AAY38791 standard; protein; 461 AA.
XX
AC AAY38791;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis antigen encoded by ORF140.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.
XX
OS Neisseria meningitidis.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB001665.
XX
PR 06-NOV-1997; 97GB-00023516.
PR 14-NOV-1997; 97GB-00024190.
PR 18-NOV-1997; 97GB-00024386.
PR 27-NOV-1997; 97GB-00025158.
PR 10-DEC-1997; 97GB-00026147.
PR 14-JAN-1998; 98GB-00000759.
PR 01-SEP-1998; 98GB-00019016.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Maignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38791.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection.
XX
PS Claim 4; Page 334; 524pp; English.
XX
CC Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis and
CC N. gonorrhoeae antigenic proteins. They are encoded by open reading
CC frames (ORFs) AA711972-Z12358. The antigenic proteins, their fragments,
CC their nucleic acids and antibodies are used for diagnosis, prevention (as
CC vaccines) or treatment of Neisseria infections, such as meningitis,
CC septicemia and gonorrhea. Both organisms are closely related. Fragments
CC of the nucleic acids are useful as hybridisation probes and antisense
CC reagents
XX
SQ Sequence 461 AA;

Query Match 47.7%; Score 51; DB 2; Length 461;
Best Local Similarity 47.4%; Pred. No. 29;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 IDNAVLPEGSLVLTGANG 22
:|:|:|:|:|:|
Db 312 VDGALAPVCSVILITGAGG 330

RESULT 14
AAY38792
ID AAY38792 standard; protein; 461 AA.

AC AAY38792;

DT 08-OCT-1999 (first entry)

DE Neisseria meningitidis strain A antigen encoded by ORF140.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.

XX Neisseria meningitidis.

OS WO9924578-A2.

XX 20-MAY-1999.

PF 09-OCT-1998; 98WO-IB001665.

XX 06-NOV-1997; 97GB-00023516.

PR 14-NOV-1997; 97GB-00024190.

PR 18-NOV-1997; 97GB-00024386.

PR 27-NOV-1997; 97GB-00025158.

PR 10-DEC-1997; 97GB-00026147.

PR 14-JAN-1998; 98GB-00000759.

PR 01-SEP-1998; 98GB-00019016.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;

XX WPI; 1999-327407/27.

DR P-PSDB; AAY38792.

XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for

XX diagnosis, treatment and prevention of infection.

XX Claim 4; Page 335; 524pp; English.

XX Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis and

XX N. gonorrhoeae antigenic proteins. They are encoded by open reading

XX frames (ORFs) AA21972-212358. The antigenic proteins, their fragments,

XX their nucleic acids and antibodies are used for diagnosis, prevention (as

XX vaccines) or treatment of Neisseria infections, such as meningitis,

XX septicemia and gonorrhea. Both organisms are closely related. Fragments

XX of the nucleic acids are useful as hybridisation probes and antisense

XX reagents

XX Sequence 461 AA;

QY Query Match 47.7%; Score 51; DB 2; Length 461;

Best Local Similarity 47.4%; Pred. No. 29;

Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 4 IDNAVLPEGSLVLTGANG 22

:|:|:|:|:|:|

312 VDGALAPVCSVILITGAGG 330

RESULT 15

AEB49401

ID AEB49401 standard; protein; 461 AA.

XX AEB49401;

XX 22-SEP-2005 (first entry)

XX N. meningitidis protein SEQ ID NO 586.

XX antibacterial; immunosuppressive; antiinflammatory; vaccine;

KW neisseria gonorrhoeae infection; gynecological; infection;

KW neisseria meningitidis infection; meningitis; neuroprotective;

XX inflammation; neurological disease; sepsis; diagnosis.

XX Neisseria meningitidis.

XX US6914131-B1.

XX 05-JUL-2005.

XX 30-APR-1999; 99US-00303518.

XX 09-OCT-1998; 98WO-IB001665.

XX (CHIR) CHIRON SRL.

XX Scarlato V, Masignani V, Rappuoli R, Pizza M, Grandi G;

XX WPI; 2005-464877/47.

XX N-PSDB; AEB49400.

XX New Neisserial nucleic acids useful for diagnosing and/or treating

XX bacterial infections, in particular meningitis and septicemia caused by

XX Neisseria meningitidis and Neisseria gonorrhoea.

XX Example 70; SEQ ID NO 586; 613pp; English.

XX The invention relates to an isolated nucleic acid molecule. The methods

XX and compositions of the present invention are useful for diagnosing

XX and/or treating Neisserial bacterial infections, in particular meningitis

XX and septicemia caused by Neisseria meningitidis and Neisseria gonorrhoea.

XX The present sequence represents the amino acid sequence of a N.

XX meningitidis protein. Note: The sequence data for this patent did not

XX form part of the printed specification, but was obtained in electronic

XX format directly from USPTO at

XX seqdata.uspto.gov/sequence.html?DocID=6914131B1.

XX Sequence 461 AA;

QY Query Match 47.7%; Score 51; DB 9; Length 461;

Best Local Similarity 47.4%; Pred. No. 29;

Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 4 IDNAVLPEGSLVLTGANG 22

:|:|:|:|:|:|

312 VDGALAPVCSVILITGAGG 330

Search completed: March 11, 2006, 05:02:25

Job time : 129.413 secs


This Page Blank (uspto)

Mon Mar 13 09:31:39 2006

us-10-616-320a-2.rat

Page 5

Search completed: March 11, 2006, 05:11:14
Job time : 31.0317 secs



This Page Blank (uspto)

```

RESULT 7
US-10-616-320A-4
; Sequence 4, Application US/10616320A
; Publication No. US20040101937A1
; GENERAL INFORMATION:
; APPLICANT: Jefferey C. Moore
; APPLICANT: Michael G. Sturr
; APPLICANT: Kathleen McLaughlin
; APPLICANT: Jaehon Kim
; TITLE OF INVENTION: PROCESS FOR REDUCING AN ALPHA-KETO ESTER
; FILE REFERENCE: 21115
; CURRENT APPLICATION NUMBER: US/10/616,320A
; CURRENT FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PRIMER
US-10-616-320A-4

```

Query Match	64.0%	Score 68.5;	DB 4;	Length 21;
Best Local Similarity	81.8%	Pred. No. 0.001;		
Matches 18; Conservative	0;	Mismatches 3;	Indels 1;	Gaps 1;

QY	MAKIDNAVLPEGSLVLTGANG	22
Db	1 MAIPDNAVL-EGSLVKVTGANG	21

```

RESULT 8
US-10-616-320A-1
; Sequence 1, Application US/10616320A
; Publication No. US20040101937A1
; GENERAL INFORMATION:
; APPLICANT: Jefferey C. Moore
; APPLICANT: Michael G. Sturr
; APPLICANT: Kathleen McLaughlin
; APPLICANT: Jaehon Kim
; TITLE OF INVENTION: PROCESS FOR REDUCING AN ALPHA-KETO ESTER
; FILE REFERENCE: 21115
; CURRENT APPLICATION NUMBER: US/10/616,320A
; CURRENT FILING DATE: 2003-07-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PRIMER
US-10-616-320A-1

```

Query Match	61.2%	Score 65.5;	DB 4;	Length 20;
Best Local Similarity	88.9%	Pred. No. 0.0029;		
Matches 16; Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1;

QY	5	DNAVLPEGSLVLTGANG	22
Db	4	DNAVL-EGSLVKVTGANG	20

RESULT 9
US-10-369-493-14080
; Sequence 14080, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

```

; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14080
;
; LENGTH: 323
;
; TYPE: PRT
;
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-14080

```

Query Match	50.5%;	Score 54;	DB 4;	Length 323;
Best Local Similarity	59.1%;	Pred. No. 5.2;		
Matches 13; Conservative	1;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      1 MAKIDNAVLPEGSVLVTGANG 22
        ||| | | | | | | |
Db      137 MALEKNGVTPQGGVLTGANG 158

```

```

RESULT 10
US-10-369-493-20309
; Sequence: 20309, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20309
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-20309

```

Query Match	45.8%	Score 49;	DB 4;	Length 339;
Best Local Similarity	64.7%;	Pred. No. 35;		
Matches 11; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;

QY	6	NAVLP	EGSLV	LTG	ANG	22
DB	164	NDALDE	GSVLV	EGTQ	180	

RESULT 11
US-09-738-626-6285
; Sequence 6285, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

! OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-39265

Query Match	44.9%	Score 48	DB 5	Length 1096
Best Local Similarity	36.8%	Pred. No. 2e+02		
Matches	7	Conservative	8	Mismatches 4
				Indels 0
				Gaps 0

```
QY      4 IDNAVLPESGLVLTGANG 22
        ::::|::|::|
Db      993 VDSGLVPEGAISLICGSAG 1011
```

```

RESULT 15
US-10-369-493-11253
; Sequence 11253, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11253
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Methanobarcina mazel
US-10-369-493-11253

```

Query Match	43.9%	Score 47;	DB 4;	Length 276;
Best Local Similarity	55.6%	Pred. No. 57;		
Matches 10;	Conservative 4;	Mismatches 4;	Indels 0;	Gaps 0;

QY 5 DNAVLPEGSLVLTGANG 22
::|||::|||
Db 112 NSAVEMEGSKIIVTGAAG 129

Search completed: March 11, 2006, 05:16:12
Job time : 102.968 secs

RESULT 7
US-11-087-099-9804
; Sequence 9804, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087, 099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9804
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Pseudomonas syringae pv. syringae B728a
US-11-087-099-9804

Query Match 41.1%; Score 44; DB 7; Length 174;
Best Local Similarity 57.1%; Pred. No. 7.5;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 5 DNALEPGLVLT 18
:|:|:|:|:|
Db 122 ENKVIDGSLVGT 135

RESULT 8
US-11-082-389-168
; Sequence 168, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131PCPN
; CURRENT APPLICATION NUMBER: US/11/082, 389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 168
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-168
Query Match 41.1%; Score 44; DB 7; Length 311;

Best Local Similarity 50.0%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 9 LPEGLVLTGANG 22
:|:|:|:|:|
Db 256 VPEGASTVITGPNG 269

RESULT 9
US-10-873-528-162
; Sequence 162, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/10/873, 528
; CURRENT FILING DATE: 2004-06-23
; PRIOR APPLICATION NUMBER: US/09/769, 787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 162
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-873-528-162

Query Match 41.1%; Score 44; DB 6; Length 363;
Best Local Similarity 47.1%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 NAVLEPGLVLTGANG 22
:|:|:|:|:|
Db 23 NITIPKGSFTLGLASG 39

RESULT 10
US-11-087-099-766
; Sequence 766, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087, 099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 766
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-11-087-099-766

Query Match 41.1%; Score 44; DB 7; Length 455;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MAKIDNAVLEPGLVLT 18
:|:|:|:|:|
Db 348 IAVILNTYIMPEGVFLIT 365

RESULT 11
US-10-793-626-284
; Sequence 284, Application US/10793626
; Publication No. US20050255478A1

```
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 284
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-284

Query Match          40.2%; Score 43; DB 6; Length 173;
Best Local Similarity 31.8%; Pred. No. 11;
Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY      1 MAKIDNAVLPEGSLLVLTGANG 22
      | :|:| :|:|:|
Db      24 MKSVTHAIPIGMMLLIIGGG 45

RESULT 12
US-11-096-568A-4407
; Sequence 4407, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4407
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(211)
; OTHER INFORMATION: Ceres Seq. ID no. 13637459
US-11-096-568A-4407

Query Match          40.2%; Score 43; DB 7; Length 211;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      6 NAVLPEGSLLVLTGANG 22
      | | :|:|:|:|
Db      12 NVSLHDGALVLTGANG 28

RESULT 13
US-11-096-568A-4406
; Sequence 4406, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4406
; LENGTH: 228
```

```
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(228)
; OTHER INFORMATION: Ceres Seq. ID no. 13637458
US-11-096-568A-4406

Query Match          40.2%; Score 43; DB 7; Length 228;
Best Local Similarity 47.1%; Pred. No. 15;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      6 NAVLPEGSLLVLTGANG 22
      | | :|:|:|:|
Db      29 NVSLHDGALVLTGANG 45

RESULT 14
US-10-467-657-3300
; Sequence 3300, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: Seqwin99, version 1.04
; SEQ ID NO 3300
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3300

Query Match          40.2%; Score 43; DB 6; Length 241;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      5 DNAVLPEGSLLVLTG 19
      ||| | :|:|:|:|
Db      58 DNAQLERKSLVMMGG 72

RESULT 15
US-11-087-099-12277
; Sequence 12277, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 12277
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Thermoplasma acidophilum
US-11-087-099-12277

Query Match          40.2%; Score 43; DB 7; Length 336;
Best Local Similarity 56.2%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      7 AVLPEGSLLVLTGANG 22
      | :|:| :|:|:|
```


Mon Mar 13 09:31:40 2006

us-10-616-320a-2.rapbn

Page 5

Db 159 AKVKEGORVLTGAGG 174

Search completed: March 11, 2006, 05:16:52
Job time : 11.873 secs

11.873 secs

This Page Blank (uspto)